
# NAVAL POSTGRADUATE SCHOOL

**MONTEREY, CALIFORNIA**

# THESIS

**OPTIMAL TRANSMITTER PLACEMENT IN WIRELESS MESH NETWORKS**

by

Paul J. Nicholas

June 2009

| | |
|---|---|
| Thesis Advisor: | David Alderson |
| Second Reader: | Brian Steckler |

**Approved for public release; distribution is unlimited**

THIS PAGE INTENTIONALLY LEFT BLANK

| REPORT DOCUMENTATION PAGE | | Form Approved OMB No. 0704-0188 |
|---|---|---|



| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | June 2009 | Master's Thesis |

| 4. TITLE AND SUBTITLE Optimal Transmitter Placement in Wireless Mesh Networks | 5. FUNDING NUMBERS |
|---|---|
| 6. AUTHOR(S) Paul J. Nicholas | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>Naval Postgraduate School<br>Monterey, CA 93943-5000 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>N/A | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER |

**11. SUPPLEMENTARY NOTES** The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| 12a. DISTRIBUTION / AVAILABILITY STATEMENT | 12b. DISTRIBUTION CODE |
|---|---|
| Approved for public release; distribution is unlimited | A |

**13. ABSTRACT (maximum 200 words)**

Wireless mesh networks are systems of wireless access points interconnected in a mesh to provide digital services to client devices via radio transmission. We consider the challenges of quickly and optimally designing a wireless mesh network. We focus on maximizing client coverage area by choice of access point locations, subject to constraints on network service, quantity and technical capabilities of access points, environmental information, and radio propagation over terrain. We create a non-differentiable, non-convex, nonlinear optimization problem to quantify the value of a given network, and use a sampling algorithm to quickly find very good solutions. We conduct field tests using commercial equipment in real-world scenarios, and conclude our technique can provide working wireless mesh network topologies.

Our techniques and associated decision support tool can be used by humanitarian assistance or disaster relief personnel and combat communications planners to quickly design wireless mesh networks. The decision support tool runs on a laptop computer, accepts map data in a generic file format, creates network topologies for virtually any type of terrain and mesh access point device, and does not require any additional software or solver licenses.

| 14. SUBJECT TERMS Wireless Mesh Networks, Humanitarian Assistance, Disaster Relief, Distributed Operations, Enhanced Company Operations, Network Design, Nonlinear Programming, Terrain Integrated Rough Earth Model, TIREM, Hata COST-231, Simultaneous Routing and Resource Allocation, Dividing Rectangles, DIRECT, Access Points, Access Point Placement, C++ | 15. NUMBER OF PAGES<br>146 |
|---|---|
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UU |

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release; distribution is unlimited**

**OPTIMAL TRANSMITTER PLACEMENT IN WIRELESS MESH NETWORKS**

Paul J. Nicholas
Captain, United States Marine Corps
B.S., United States Naval Academy, 2003

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN OPERATIONS RESEARCH**

from the

**NAVAL POSTGRADUATE SCHOOL**
**June 2009**

Author: Paul J. Nicholas

Approved by: David Alderson
Thesis Advisor

Brian Steckler
Second Reader

Robert Dell
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

# ABSTRACT

Wireless mesh networks are systems of wireless access points interconnected in a mesh to provide digital services to client devices via radio transmission. We consider the challenges of quickly and optimally designing a wireless mesh network. We focus on maximizing client coverage area by choice of access point locations, subject to constraints on network service, quantity and technical capabilities of access points, environmental information, and radio propagation over terrain. We create a non-differentiable, non-convex, nonlinear optimization problem to quantify the value of a given network, and use a sampling algorithm to quickly find very good solutions. We conduct field tests using commercial equipment in real-world scenarios, and conclude our technique can provide working wireless mesh network topologies.

Our techniques and associated decision support tool can be used by humanitarian assistance or disaster relief personnel and combat communications planners to quickly design wireless mesh networks. The decision support tool runs on a laptop computer, accepts map data in a generic file format, creates network topologies for virtually any type of terrain and mesh access point device, and does not require any additional software or solver licenses.

THIS PAGE INTENTIONALLY LEFT BLANK

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

# LIST OF ACRONYMS AND ABBREVIATIONS

ACO – ant colony optimization

AM – amplitude modulation

AP – access point

BER – bit error rate

bps – bits per second

BPSK – bi-phase shift keying

COTM – communications-on-the-move

DES – discrete event simulation

DHS – Department of Homeland Security

DIRECT – DIviding RECTangles

DLL – dynamic link library

DOD – Department of Defense

DR – disaster relief

DSSS – direct sequence spread spectrum

DTED – digital elevation terrain data

ECO – Enhanced Company Operations

FDMA – frequency division multiple access

FHSS – frequency hopping spread spectrum

FM – frequency modulation

GA – genetic algorithm

GPS – Global Positioning System

GUI – graphical user interface

HA – humanitarian assistance

kbps – kilobits per second (1,000 bits per second)

LOS – line of site

MA – multiple access

MIP – mixed integer program

MOA – multiple objective algorithm

OFDM – orthogonal frequency division modulation

PTP – point-to-point

QAM – quadrature amplitude modulation

QoS – quality of service

QPSK – quadrature phase shift keying

RCA – radio coverage analysis

RSS – received signal strength

SLA – service level agreement

SNR – signal to noise ratio

SPEED – System Planning, Engineering and Evaluation Device

SRRA – Simultaneous Routing and Resource Allocation

SRRA+C – Simultaneous Routing and Resource Allocation with Coverage

TDMA – time division multiple access

TIREM – Terrain Integrated Rough Earth Model

USGS – United States Geological Survey

USMC – United States Marine Corps

UTM – Universal Transverse Mercator

WMN – wireless mesh network

# EXECUTIVE SUMMARY

Wireless mesh networks (WMNs) are systems of wireless access points (APs) interconnected in a mesh to provide digital services to client devices via radio transmission (Zhang et al., 2006, pp. 564–565). Military and civilian organizations can benefit from the advantages provided by WMNs. During combat operations, WMNs can quickly and securely relay time-critical information such as intelligence reports, tactical orders, and location sensor readings to separated small units. During humanitarian assistance (HA), and disaster relief (DR) operations, WMNs can provide maps, floor plans, video surveillance, emergency aid requests, and other critical information to personnel. The ability of WMNs to reroute traffic dynamically when APs are lost or added to the network, and the ability to operate with no infrastructure other than a local power source (such as a battery or small generator) makes them particularly well-suited to austere environments.

The physical topology (i.e., the locations of the wireless APs) of a WMN critically affects its performance. Wireless APs must be placed and configured to provide service to clients in desired areas, while meeting restrictions on quantity, placement, and characteristics of APs, and requirements for redundancy, bandwidth, and other service standards. Consideration must also be given to the effects of terrain and other aspects of the operating environment on radio wave propagation. Since the aforementioned applications are highly time-sensitive, the network must be designed quickly and with as little guesswork as possible.

We consider the challenges of quickly and optimally designing a wireless mesh network. We focus on maximizing client coverage area by choice of access point locations, subject to constraints on network flow and power allocation, number and technical capabilities of APs, environmental information, and radio propagation over terrain.

The imperative to maximize client coverage creates an incentive to place APs farther away from one another. This will reduce or eliminate redundant coverage,

maximizing the area receiving adequate client service from a wireless AP. However, the constraints of network service and AP capabilities will create an incentive to place APs nearer each other. This will increase AP-to-AP network performance and ensure technical constraints are satisfied. The tension between these competing design goals is the core of our optimization problem.

We decompose the problem into two major subproblems. The first calculates client coverage using the Terrain Integrated Rough Earth Model (TIREM) (Alion Science and Technology Corporation, 2007), given AP locations, operating characteristics, and terrain and environment information. The second subproblem determines the optimal routing and power allocation scheme to quantify the value of network flow, solved using the Simultaneous Routing and Resource Allocation (SRRA) techniques of Xiao et al., (2004). After combining the SRRA and coverage subproblems via a penalty function, we name our overall formulation SRRA+C. SRRA+C is a non-differentiable, non-convex, nonlinear optimization problem. We utilize the DIviding RECTangles (DIRECT) algorithm of Jones et al., (1993) to iteratively sample the SRRA+C objective function in pursuit of an optimal solution.

We validate our theoretical results by comparing them to an exhaustive enumeration of the solution space, and find the DIRECT algorithm determines excellent solutions to the SRRA+C problem very quickly. We also conduct field tests using commercial equipment in real-world scenarios, and conclude the SRRA+C formulation can provide working WMN topologies.

To our knowledge, we are the first to formulate and solve a WMN design problem to maximize client coverage, given constraints on network service, quantity and characteristics of AP devices, environmental information, and radio propagation over terrain. Our techniques and associated decision support tool can be used by HA/DR personnel and combat communications planners to quickly design WMNs to support their specific operations. The decision support tool runs on a laptop computer, accepts map data in a generic file format that is widely available on the Internet, creates network topologies for virtually any type of terrain and mesh AP device, and does not require any additional software or solver licenses.

# ACKNOWLEDGMENTS

I am deeply indebted to my advisor, Dr. David Alderson. This thesis would not have been possible without your guidance and insight. The latitude you afforded me made this last year an incredibly rich learning experience, and your ability to quickly discern and parse complex problems was humbling to behold. It was my honor to work with you; I hope to do so again in the future.

Thank you to my technical reviewers, Brian Steckler, Johannes Royset and Rex Buddenberg of the Naval Postgraduate School. I appreciate your time and candor. I owe a special debt of gratitude to Ramsey Meyer of the Hastily-Formed Networks Research Group, without whom field testing would not have been possible. Thank you for numerous hours working in the badlands of Fort Ord.

I must thank Ray Trechter and David Duggan of Sandia National Labs for serving as a sounding-board for ideas, and providing the Zen-like alcove where I was able to accomplish much of this work. The mere presence of so many great minds at Sandia surely had a positive impact on my research.

I am indebted to so many individuals who graciously provided their insight or helped me overcome snags. This list is certainly not exhaustive: LtCols Serg Posadas and Carl Oros, Carlos Borges, Jerry Brown, Matt Carlyle, Michael Clement, Debbi Kreider and Glen Koue of the Naval Postgraduate School; Karen Griego-Peralta at SNL; Maj Jian Xu and Brian Mask at Marine Corps Systems Command; Ken Wilson at Marine Corps Tactical Systems Support Activity; Doug Galarus, Bill Jameson and Jeff Sharkey at Montana State University; Carey Beall at the JSC; Josephine Cantalupo at Alion; D.J. Shyy at Mitre; Luke Berndt and Trent DePersia at DHS; Pat Toole at Toyon; Joel Eichelberger at New Mexico State University; and Amarildo Vieira at Motorola.

I would be lost, cold and hungry without my wife and best friend, Melissa. Thank you for your love, patience, unwavering support, and ability to turn each of my many spectacular failures into the butt of a joke. Thank you especially for keeping that small human occupied while her Daddy spent countless hours punching keys at the computer.

THIS PAGE INTENTIONALLY LEFT BLANK

# I. INTRODUCTION

## A. WIRELESS MESH NETWORKS: PROPERTIES AND USAGE

Wireless mesh networks (WMNs) are systems of wireless *access points* (APs) interconnected in a mesh to provide digital services to client devices via radio transmission. These client devices can be laptops, personal digital assistants (PDAs), sensor systems, or other electronic devices requiring network connectivity (Zhang et al., 2006, pp. 564–565). WMNs support any type of digital communication, including email, Internet access, file transfer, and voice communication, and they can transmit this information rapidly, reliably, and efficiently.

WMNs can be quickly deployed and configured to fulfill communications requirements, and they have several characteristics that make them particularly well-suited to austere environments (i.e., scenarios where communications infrastructure is seriously degraded, destroyed or non-existent). This includes the ability to reroute traffic dynamically when APs are lost or added to the network and the ability to operate with no infrastructure other than a local power source (such as a battery or small generator). Unlike devices connected to wireless local area networks (WLANs) or "hotspots," clients on WMNs are free to roam from the coverage area of one wireless AP to another without loss of service.

Military and civilian organizations can benefit from the advantages provided by WMNs. The Enhanced Company Operations (ECO) concept of the United States Marine Corps (USMC) relies on separation and coordination of tactical units to gain an advantage over the enemy. This concept requires "voice, data, and surveillance fused into a single common operating picture… Tactical units must gravitate from push-to-talk radio systems to mobile ad hoc mesh networking" (USMC, 2008, p. 5). WMNs can quickly and securely relay time-critical information such as intelligence reports, tactical orders, and location sensor readings to separated small units in support of ECO.

The Department of Homeland Security (DHS) requires that the public safety community have "the ability to communicate and share information as authorized when it is needed, where it is needed, and in a mode or form that allows the practitioners to effectively use it." DHS also states, "Data communications are becoming increasingly important to public safety practitioners to provide the information needed to carry out their missions" (DHS, 2006, pp. 1, 3). During humanitarian assistance and disaster relief (HA/DR) operations, WMNs can provide maps, floor plans, video surveillance, emergency aid requests, and other critical information to personnel.

The physical topology (i.e., the locations of the wireless APs) of a WMN is critical to its performance. Network engineers must place and configure wireless APs to provide service to clients in desired areas, while meeting restrictions on number, placement, and characteristics of APs, as well as requirements for coverage, bandwidth, and other service standards. Consideration must also be given to the effects of terrain and other aspects of the operating environment on radio wave propagation. Because combat and HA/DR operations are highly time-sensitive, the WMN must be designed quickly and with as little guesswork as possible.

## B. PROPERTIES OF WIRELESS MESH NETWORKS

A wireless mesh network consists of fixed APs that use separate radio systems to provide two levels of network connectivity (we use the term *access point* generically for any type of wireless mesh transmitter or base station device). The fixed position of APs differentiates WMNs from so-called "ad hoc" networks, where APs can be constantly moving (Zhang et al., 2006, p. 565). The first connectivity level of a WMN supports AP-to-AP communication, providing the *backhaul radio network* that routes traffic from source to destination. The second level supports AP-to-client radio communication (see Figure 1, below). Client devices, such as laptops, PDAs, sensors, or other electronic devices, can be mobile within the WMN *client coverage area*. While a client device is roaming within a contiguous coverage area, APs may seamlessly "hand off" the client to a more appropriate AP without loss of client network connectivity. WMNs may connect to an outside network or the Internet through a gateway (including via a satellite uplink).

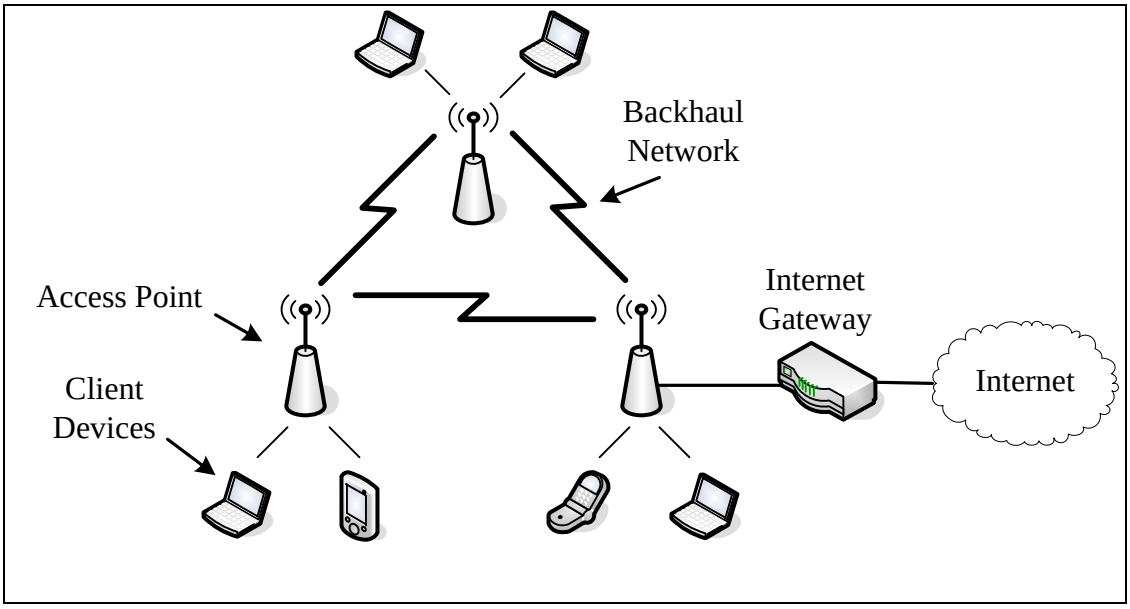


***Figure 1: Depiction of a wireless mesh network***

WMNs have two important properties that make them particularly useful for time-critical deployments such as combat and HA/DR operations. WMN APs, once placed, can "self-form" a network with little or no oversight. AP devices detect the presence of other nearby APs, and automatically determine the most efficient method of passing traffic within the mesh. Similarly, WMNs can determine when an AP has been removed from the network for any reason, and will automatically adapt or "self-heal" to reroute traffic.

Additionally, WMNs are able to adapt to changing client demand and quality-of-service (QoS) constraints. Should a particular AP's coverage region contain more clients or a client requiring more throughput (such as during a video teleconference), the WMN will reroute traffic to best satisfy demand. Note the determination of "best" is often dictated by proprietary routing algorithms.

Figure 2 shows the three different types of WMN architecture. In the *infrastructure mesh* type of WMN, client devices communicate only to a single AP at a time, and hence do not route traffic. In a *client mesh* architecture, there are no dedicated

APs: each client device serves as both a user terminal and as an AP, routing traffic directly to other client devices. A *hybrid mesh* architecture consists of both traditional APs and client devices capable of serving as APs (Zhang et al., 2006, pp. 564–567). This thesis considers only infrastructure mesh architectures in which wireless APs alone provide traffic routing services; clients do not perform any routing. Additionally, we assume that each AP, once positioned, will remain stationary.

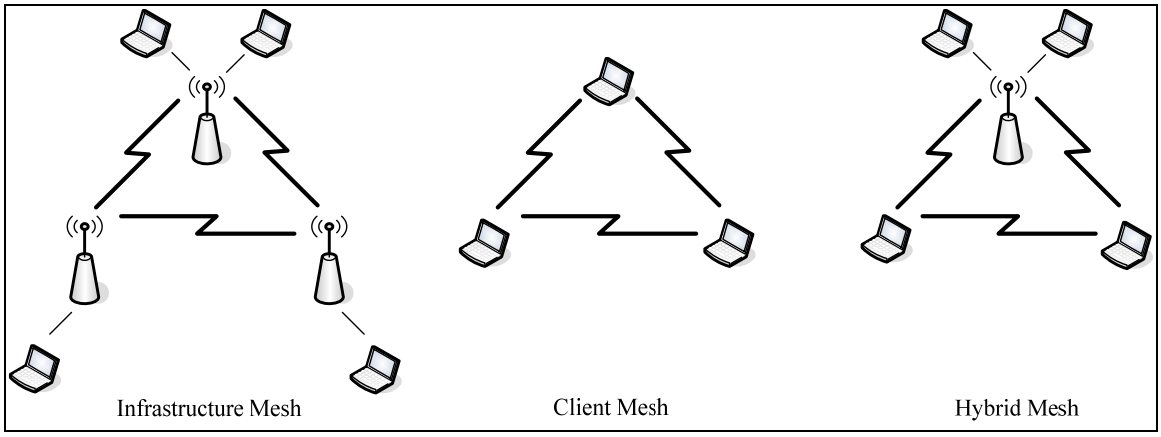


***Figure 2: Depictions of the three types of WMNs***

## C. RESEARCH PROBLEM STATEMENT

We consider the task of a communications officer or field technician who must quickly decide where to place APs so that the resulting WMN has good performance. While the value of a WMN can be measured in many ways (e.g., throughput, power usage, financial cost, security, availability or reliability), we focus on maximizing client coverage area subject to constraints on network flow.

The communications officer or field technician faces a fundamental tension when designing a WMN. The imperative to maximize client coverage creates an incentive to spread APs far apart. However, the need for good backhaul communication among APs creates an incentive to keep them closer together. Further complicating the design

problem are the technical capabilities of a limited number of APs, background and environmental noise, and the complexities of radio propagation over real terrain.

Given real terrain data and a desired client coverage area, we evaluate the performance of proposed AP locations by solving two major subproblems. The first calculates client coverage using a point-to-point path loss propagation model, given AP locations, operating characteristics, terrain data and environment information. The second subproblem determines the optimal routing and power allocation scheme for the backhaul network, solved using the Simultaneous Routing and Resource Allocation (SRRA) techniques of Xiao et al., (2004). We combine the SRRA and coverage subproblems via a penalty function to obtain SRRA+C, a non-differentiable, non-convex, nonlinear optimization problem. We utilize the DIviding RECTangles (DIRECT) algorithm of Jones et al., (1993) to iteratively sample the SRRA+C objective function in pursuit of an optimal solution.

We create a graphical decision support tool to solve the SRRA+C problem using Microsoft Visual C++ (Microsoft Corporation, 2009). The stand-alone tool reads and then graphically displays digital terrain elevation information, obtains its best solution to the SRRA+C problem, displays the resulting network and coverage regions, and calculates detailed point-to-point propagation information for any two given points.

We validate the quality of our results by comparing them to an exhaustive enumeration of the solution space, and we find the DIRECT algorithm determines excellent solutions to the SRRA+C problem very quickly. We also conduct field tests using commercial equipment, and conclude that WMNs generated from our SRRA+C formulation function in real-world scenarios as predicted.

To our knowledge, this is the first technique to determine WMN AP locations given constraints on number and technical characteristics of APs, targeted coverage area, service requirements, and the operating environment using an algorithm with guaranteed convergence. Our model and its solution can help decision-makers to quickly design

WMNs in support of combat or HA/DR operations, and we can integrate these contributions into existing WMN planning tools to quickly provide useful design information.

## D. LITERATURE REVIEW OF PREVIOUS WORK

Previous academic work has developed techniques to design mesh networks using various algorithms and heuristics. None of the existing approaches determine AP locations using an algorithm with proven global convergence while satisfying constraints of number and characteristics of APs, service requirements, and placement restrictions of APs while also utilizing detailed terrain data.

Calegari et al., (2001) employ three different types of optimization algorithms to find the minimum number of transceivers to provide "good" service coverage (around 90%) to a targeted region. Their work utilizes a radio propagation tool to determine actual site coverage in areas of France and Switzerland, given sets of no more than 150 candidate transceiver sites. Their results compare the performance of greedy, Darwinism, and genetic algorithms in solving the network design problem.

Calegari et al.'s greedy algorithm successively adds single APs to the solution set to maximally increase service coverage during each iteration. This algorithm is by far the fastest of the three (taking just a few seconds), and provides a very good solution requiring 58 transceivers. Unfortunately, the greedy algorithm is prone to falling into inferior local optima.

The Darwinism algorithm iteratively guesses the number of APs needed, then computes the number of shared (overlapping) cells to determine performance. This algorithm not only provided the worst results (requiring 82 transceivers), but also took about two minutes to execute.

The genetic algorithm (GA) assigns possible solutions as chromosomes of individuals in a randomly-created population. Only the fittest individuals are allowed to survive to the next generation, with fitness being determined by the percentage of targeted service area covered by the solution set chromosome. Pairs of individuals mate, create offspring, and mutate using probabilistic processes. This algorithm is

computationally the most expensive (requiring 24 minutes to execute), and produced marginal results of 70 required transceiver locations.

Calegari et al. also investigate a variation of the genetic approach. This technique divides the population among separate "islands," and allows the fittest individual on an island to replace a probabilistically-selected individual on a nearby island. This approach greatly reduces the computational workload (requiring 4 minutes and 30 seconds to execute), and provides the best solution of 57 transceivers. The technique is also much less prone to fall into bad local optima than the greedy algorithm. As the concept of separate islands lends itself to parallel processing, Calegari et al. show that by running this algorithm on ten parallel computers, processing time is 7.8 times faster, and will execute in 12 seconds on 40 parallel computers.

Allen et al., (2002) design cellular networks at minimum cost while considering constraints of coverage, traffic capacity, interference, and hand-over between cellular coverage regions. They use simulated annealing (SA) iteratively to create candidate solutions by perturbing an existing solution. A candidate solution is selected for the next iteration if it produces a better objective value than the original, or probabilistically based on a variable denoted temperature, and a cooling schedule. As the process continues, the cooling schedule gradually lowers the temperature, i.e., the probability of choosing a less-desirable candidate solution for the next iteration gradually decreases as a function of the cooling schedule. Using a realistic-sized network of 40x170 kilometers and nearly 30,000 coverage test cells, the SA algorithm produces very good results in 119 minutes on a 550 MHz personal computer.

Raisanen and Whitaker (2005) present a novel approach using a multiple objective algorithm (MOA) to provide the greatest service coverage at the least financial cost. Their method provides a range of possible solutions that approximates a Pareto frontier between these two competing objectives. They compare four different genetic algorithms to solve the multiple objective function: the Strength Pareto Evolutionary Algorithm version II (SPEA2) (Zitzler et al., 2001), the Non-Dominated Sorting Genetic

Algorithm (NSGA) (Deb et al., 2000), the Pareto Envelope-based Selection Algorithm (PESA) (Corne et al., 2000), and the Simple Evolutionary Algorithm for Multiobjective Optimization (SEAMO) (Valenzuela, 2002).

Using realistic-sized problems (up to 45x45 kilometers, with 244 candidate locations), they evaluate these four GAs based on obtained objective values, speed of convergence, average population spread, and average speed of execution. This work assumes isotropic radiation patterns, and does not consider terrain information. The NSGA-II and SPEA2 GAs performed by far the best, with PESA and SEAMO performing considerably worse. Execution time for the slowest GA on the largest problem took roughly 50 minutes.

Beljadid et al., (2007) construct a mixed integer linear program (MIP) to design WMNs that meet location, power, and routing constraints while minimizing cost. The model assumes line-of-site connectivity between APs and does not consider terrain information. Even without these added complexities, the MIP is NP-hard and can only be solved exactly for very small idealized networks. Bhatnagar et al., (2006) show similar results: networks of 15 nodes take several days to solve exactly.

Amaldi et al., (2008) construct three MIPs to determine the number and locations of wireless APs that meet restrictions on channel assignment, interference, and traffic routing. They introduce two heuristics to solve the more computationally-expensive MIPs to within 5% of the optimum solution. The first heuristic relaxes several integer variables to continuous variables (making the problem computationally easier), and then examines neighboring solutions by iteratively adding or deleting wireless APs. The second heuristic iteratively examines solutions by considering interference and traffic load. Amaldi et al., (2008) also use discrete event simulation (DES) to simulate the predicted traffic flows of one of their models. This work does not consider terrain data, and assumes propagation patterns are given.

Sharkey (2008) utilizes ant colony optimization (ACO) to minimize financial cost of radio networks to support rural transportation systems. His algorithm determines the location of radio towers to meet restrictions on coverage, bandwidth, maximum delay, and redundancy constraints, while utilizing digitized terrain information from the United

States Geological Survey (USGS). Ants are placed at terminal (coverage nodes) of a network represented using a minimum Steiner tree. Each ant, in a randomly selected order, then moves toward the global gateway node. The ants make path decisions based on the cost of edges and the strength of pheromones left by previous ants. Pheromones are updated after each iteration based on total path cost. All ants from the same node share a tabu list of edges already visited, causing each to choose a new path. The ACO algorithm provides considerably better results than a 2-approximation algorithm for realistic-sized problems.

He et al., (2004) utilize the DIviding RECTangles (DIRECT) algorithm of Jones et al., (1993) to find good AP placements in indoor wireless networks. They use highly accurate (and computationally expensive) ray-tracing techniques to predict radio propagation within a building, creating what they dub a site-specific system simulator for wireless system design ($S^4W$). We build on the techniques of He et al., (2004) to formulate our concept of coverage area, and utilize the DIRECT algorithm in a similar fashion.

Our work also builds upon the Simultaneous Routing and Resource Allocation (SRRA) methodology of Xiao et al., (2004). Their dual decomposition technique solves the optimal routing and resource allocation scheme of a wireless network. That is, given AP positions and capabilities, and information about the operating environment, Xiao et al., (2004) show how to optimally route network traffic and allocate AP resources (such as transmission power) to satisfy several types of objective functions. We utilize a similar framework to calculate network performance within our optimization problem.

Shankar (2008) also utilizes the SRRA techniques of Xiao et al., (2004) to determine optimal jammer placement in WMNs. He uses the SRRA formulation to determine network flow among fixed nodes, and then adopts the attacker-defender formulation techniques of Brown et al., (2006) to calculate jammer locations among a known set of possible positions to maximally damage this flow. Our work is a natural complement to Shankar's research, addressing the question of how to design an initial WMN.

## E. STRUCTURE OF THESIS AND CHAPTER OUTLINE

Chapter II provides a more detailed description of WMNs, and presents some of the challenges and techniques of network design. In Chapter III, we describe our problem element by element, gradually building to the overall SRRA+C formulation and the techniques we use to solve it, including the DIRECT algorithm. In Chapter IV, we validate each aspect of our formulation, and provide detailed performance results of DIRECT in solving SRRA+C. We also provide the results of field-testing conducted using our decision support tool. Chapter V summarizes our results and provides suggestions for follow-on research.

# II. DESIGNING WIRELESS MESH NETWORKS

Myriad factors affect the design of wireless mesh networks; it is beyond the scope of this thesis to describe them all. We present a brief overview of the more critical design issues, including all those considered in our problem formulation. We recommend Olexa (2005) and Zhang et al., (2006) to the reader who desires more information on WMN design, and Balanis (2005) for a very detailed text on antenna theory.

## A. WMN DESIGN CONSIDERATIONS

Because WMNs are able to self-form, self-heal, and reroute traffic automatically based on QoS and client demand, *network performance is ultimately a function of AP placement*. WMN engineers must determine AP locations to best meet their specific design goals. We detail some of the aspects considered in determining AP location.

### 1. Access Point and Client Device Characteristics

#### *a. Static Characteristics*

We assume the network engineer selects the following AP and client device characteristics prior to network employment, and that these aspects do not change while the WMN is in use:

(1) Antenna Height. To an extent and depending on transmission frequency, the higher an antenna, the farther an AP can propagate its radio signal, since most obstructions in the path of the radio waves are close to the ground. For this reason, WMN APs in urban environments are often placed on telephone or traffic signal poles. However, a transmitting antenna that is too high may be out of range from a client device, even if no object obstructs the path. For example, a small home AP placed atop a very tall building is unable to connect to a laptop three miles away.

(2) Antenna Gain. The gain of antennae is one of the most critical aspects in designing WMNs. The concept of gain incorporates both efficiency and direction, and is generally measured in reference to an isotropic transmitter (Balanis,

2005, pp. 58–59). An isotropic transmitter radiates equally in all directions, in a spherical shape. For the purposes of illustration, a simple light bulb can be considered an isotropic transmitter (see left side of Figure 3, below). By placing a cone-shaped reflector on one side of the light bulb, much more light is directed in one direction, and much less in the opposite (right side of Figure 3). However, the same total amount of light is leaving the bulb itself. In this case, the cone and light bulb system is exhibiting high gain in the direction where light is being cast, and low gain elsewhere.

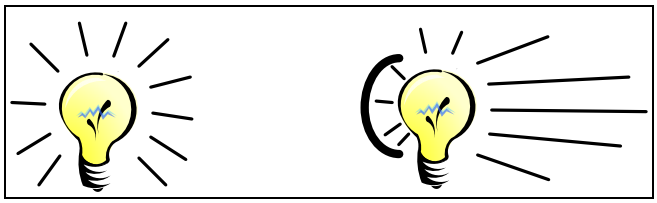

***Figure 3: Analogy of an isotropic antenna (left) and high gain antenna***

The gain of an antenna is the ratio of radiation intensity in a particular direction to the radiation intensity in that direction using an isotropic radiation pattern. Gain is generally measured in decibels (dB) or decibels relative to an isotropic transmitter (dBi) (Balanis, 2005, p. 60). The gain stated for a particular antenna is usually in the direction of the strongest lobe in its particular radiation pattern. The gain of different types of antennae can be leveraged when designing WMNs: high-gain antennae can be used for backhaul connections between APs, and low-gain omnidirectional antennae can be used for providing client coverage (Olexa, 2005, pp. 39–40).

By the property of reciprocity, in general the radiation pattern of an antenna is equivalent to its receiving pattern (Balanis, 2005, pp. 127–132). That is, an antenna with high gain in a particular direction transmits stronger signals and is more sensitive to received signals in that direction. Hence, a high-gain antenna increases received signal strength in the same direction and magnitude as transmitted signal strength. Continuing the cone and light bulb example, the surface of the light bulb receives more incoming light from the direction opposite the cone.

Because total radiation power is conserved, a high gain antenna is also very directional. A flashlight with a focusable lens is an analogy of this phenomenon. By focusing the beam, the light pattern is more tightly concentrated, but the beam must be pointed with more precision. The directionality of a high gain antenna increases the number of variables in designing a WMN, as the horizontal and vertical (tilt) direction of each antenna must be determined prior to network deployment. In this thesis, we assume all antennae are omnidirectional, and do not consider direction or tilt.

(3) Polarization. The polarization of an antenna describes the orientation of its electromagnetic field in relation to the horizon of the earth. Antennae can be polarized horizontally, vertically, or circularly (constantly changing). It is particularly important in line-of-site (LOS) communications (which WMNs can utilize) to ensure the transmitting and receiving antennae are similarly polarized (Olexa, 2005, p. 153).

(4) Transmitter Losses. Within an electronic device, signals are degraded by the media through which they flow, including cables, connectors, and other electronics. These losses can be minimized by reducing the distance the signal must travel prior to being transmitted or processed, and by using high quality, low impedance equipment (Balanis, 2005, pp. 873–874).

(5) Transmission Frequency and Bandwidth. Frequency is the number of cycles an electromagnetic wave passes through per second, measured in hertz. Bandwidth is the range of frequencies a device can transmit. Bandwidth and frequency determine the maximum amount of information that can be transmitted in a given amount of time. Larger bandwidths are capable of carrying more information. In general, higher frequencies can carry more information, but are more susceptible to attenuation during transmission and hence are useful over shorter distances than lower frequencies (Olexa, 2005, p. 69).

(6) Modulation Scheme. To impress information upon an electromagnetic signal, the signal must be modified in some way to represent the information. The most common methods of modulation are amplitude modulation (AM) and frequency modulation (FM). As their names imply, amplitude and frequency

modulation respectively vary the amplitude (power) and frequency of a signal over time. More complicated modulation schemes include Bi-Phase Shift Keying (BPSK), Quadrature Phase Shift Keying (QPSK), and Quadrature Amplitude Modulation (QAM). In general, more complex modulation schemes can carry more information, but require a stronger signal and more complex electronics (Olexa, 2005, pp. 48–49).

(7) Multiple Access Techniques. In a simple half-duplex or "walkie-talkie" network, a radio transmits and then waits while the distant side transmits. Duplexing combines signals to allow continuous two-way (full duplex) communication, where each side is able to essentially transmit and receive at the same time. Duplexing is accomplished by either dividing the available frequencies within the operating bandwidth, or dividing the amount of time each device is able to transmit.

Multiple Access (MA) is an extension of the full duplex idea. Rather than just two devices, MA enables multiple devices to communicate concurrently. Frequency Division Multiple Access (FDMA) divides the available bandwidth into channels, and assigns a channel to devices that wish to communicate. Each device is able to use its assigned channel continuously (see left side of Figure 4, below). Time Division Multiple Access (TDMA) similarly assigns a "time slot" to each device during which that device is allowed full use of the available bandwidth. By quickly cycling through all time slots, TDMA-enabled devices can essentially operate continuously (right side of Figure 4). Other types of MA techniques included Frequency Hoping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS), and Orthogonal Frequency Division Modulation (OFDM). Utilization of any MA technique involves tradeoffs of capacity, complexity, and cost (Olexa, 2005, pp. 48–49, 52–54).

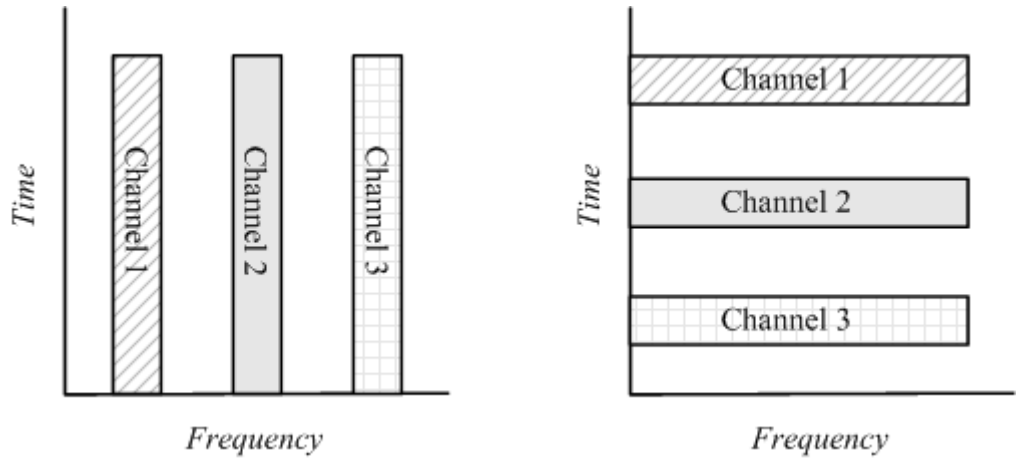


***Figure 4: Illustrations of frequency division (left) and time division multiple access schemes***

### *b. Dynamic Characteristics*

An operating WMN determines traffic routing decisions in real time. More specifically, WMN APs must calculate how to route network traffic from source to destination while satisfying physical resource limits (such as maximum available transmit power) that determine how much and where traffic can be sent. Hence, routing and allocation of available resources at an AP are co-dependent (Xiao et al., 2003). The protocols used to make these decisions are often proprietary and vary widely by manufacturer and device. Some of the characteristics relevant to these decisions include the following:

(1) Client and QoS Demand. The number of clients and the type of connectivity each uses can vary widely within a WMN. If one particular area is experiencing great demand, APs can adjust routing schemes dynamically to spread the burden throughout the network. Likewise, if a particular client has a higher QoS requirement (such as during a video teleconference), it is possible to create a dedicated path through the network.

(2) Efficiency. APs can route traffic along the least congested path to the destination. The congestion of any particular path is a function of the current

traffic flow, number of hops, required routing overhead (the additional workload required to send traffic through a certain path), and other factors.

(3) Near-Far Effect. Consider two client devices communicating with an AP at different distances. All other things being equal, the client device that is closer to the AP receives a stronger signal and hence better service than the distant client, due to the nature of radio propagation physics. This phenomenon is known as the near-far effect. Several manufacturers have proprietary protocols to reduce its effect and more evenly distribute service among clients.

## 2. Environment Characteristics

For a receiver to communicate with a distant transmitter, a radio signal of sufficient strength must be received. Once a radio wave has been emitted from the surface of an antenna, the physics of the wave in the operating environment determine how it propagates and how strong it is at the receiving antenna. A full explanation of radio physics is beyond the scope of this thesis, but we consider several important concepts.

### *a. Terrain and Land Use*

In an ideal situation, where there are no obstructions between a transmitter and receiver, the signal path is strictly line-of-sight (LOS). The only loss of signal strength is due to *free space loss*, which occurs because the radio waves gradually widen and thus diffuse over distance (Olexa, 2005, pp. 69–70) (see Figure 5 below).

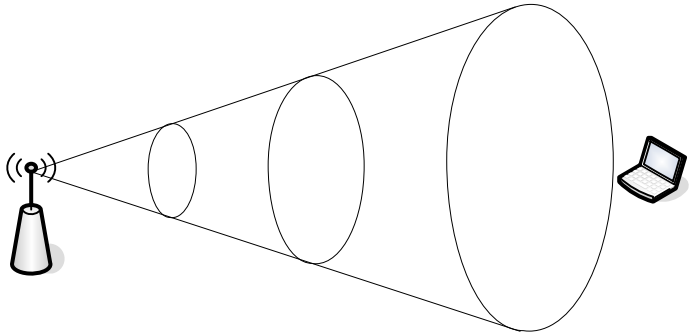

***Figure 5: Illustration of wave diffusion as a function of distance***

In the real world, however, there can be many obstructions in the path of the signal. Obstructions such as buildings and terrain can partially or completely absorb a signal before it reaches its intended destination. Smaller objects, such as tree leaves or a passing car, can reflect or partially absorb radio signals. This can cause fading of the entire signal, or selective fading of certain frequencies in the signal (Olexa, 2005, pp. 69, 73).

Even obstructions not directly in the LOS path can negatively affect propagation. Fresnel zones are a continuous series of imaginary ellipsoids between transmitter and receiver. The diameter of each Fresnel zone is a function of signal frequency and distance from the transmitter. If an object protrudes significantly into the zone, radio waves can bounce off the object and be received out-of-phase from the rest of the transmission, potentially degrading or even completely cancelling the signal (see Figure 6 below) (Olexa, 2005, pp. 69–71).

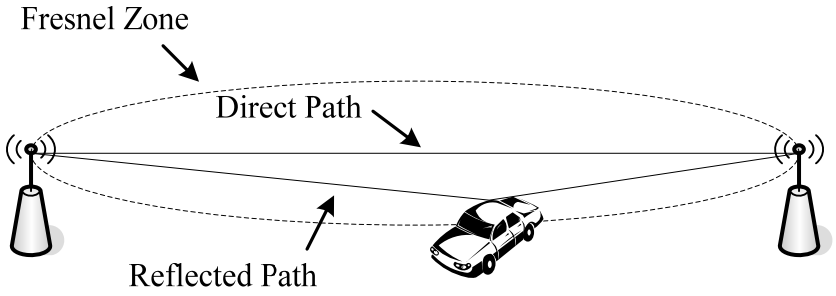


***Figure 6: Illustration of the effect of an object in the Fresnel zone***

Fading and Fresnel zone obstructions are collectively known as multipath interference. The effects can be so localized that moving a receiver just inches can completely change signal quality (Olexa, 2005, p. 72). Also, different operating frequencies behave very differently: higher frequency waves are generally more easily absorbed, while lower frequency waves can be reflected off the earth's atmosphere to increase their range.

### *b. Other Characteristics*

Environmental factors such as humidity, temperature, and surface refractivity and permittivity can affect how radio waves propagate. Background noise—both man-made and naturally occurring—can also affect the quality of the received information. A strong signal may be received from a transmitter, but high background noise can essentially drown it.

Engineers incorporate a fade margin into network design to account for the huge uncertainties of propagation. Fade margin is essentially a "slop factor" used to compensate for the cumulative effects of all signal losses not otherwise estimated. The use of a high fade margin when constructing a WMN leads to a conservative topology that is more likely to avoid being overcome by signal loss.

## 3. Gauging Network Performance

Clearly, network engineers must consider many important radio propagation properties. The interaction of these properties, along with system cost and available manpower, often create competing demands when designing a WMN. Systems operating at high frequencies can carry more traffic, but may work only over short distances. Quadrature amplitude modulation devices can process more information, but are more expensive. Network performance metrics can assist in making decisions between such tradeoffs, and quantitatively evaluating the performance of a network. The following are some of the available network performance metrics.

### *a. Cost and Profit*

Cost and profit are perhaps the most straightforward methods of calculating network performance. Cost can be calculated by summing the costs of equipment, maintenance, operation, and other activities for the network, and is generally measured in money or labor. This metric places a premium on using the simplest, least-cost technology to satisfy other design constraints. Beljadid et al., (2007), Allen et al.,

(2002), and Sharkey (2008) minimize cost in their respective research. Raisanen and Whitaker (2005) balance cost and service coverage.

The monetary profit generated by a network can be derived from customer use. A premium would be placed on ensuring clients with high bandwidth needs (i.e., those who are paying the most) have access to it. Profit is also a metric in a network based on a contract containing service level agreements (SLAs) tied to any other type of performance metric.

### *b. Power Consumption*

Power consumption is useful as a metric in situations where power is scarce or expensive. For instance, a network designed to support disaster-relief operations may require portable generators or batteries as power sources for both APs and client devices. Minimizing the use of power enables the network to be operated for longer periods of time and with less fuel or fewer replacement batteries. A network designed to minimize consumption favors both WMN devices that consume little power, and propagation paths that do not require large amounts of transmission power to achieve desired performance levels (such as short, unobstructed paths).

### *c. Throughput*

There are many methods to compute network throughput, in both the academic and manufacturer literature. These include measuring the maximum, minimum, or average throughput from any client to the network gateway, or to any other client. Another technique measures the average utilization rate of all links within the network. Still another simply sums the total flow of traffic through the gateway node over a specified period of time.

Higher throughput can be achieved by transmitting more power or at a higher frequency, or using a larger bandwidth or a more complex modulation scheme. It can also be achieved by placing APs in locations where their transmitted signals are

better received, such as in closer proximity to each other, in areas less affected by path loss such as atop hills, buildings, or posts, and in areas less affected by noise, such as away from other electromagnetic fields.

Network throughput can also be a performance constraint. Beljadid et al., (2007), Amaldi et al., (2008), Allen et al., (2002), Sharkey (2008), and this thesis all use some measure of throughput as a constraint. We employ the maximum utility Simultaneous Routing and Resource Allocation (SRRA) concept of Xiao et al., (2004) to measure network throughput. This technique is explained in greater detail in Chapter III.

#### *d. Availability and Reliability*

A network, measured by availability, places a premium on ensuring clients are able to access the network a high percentage of the time. In a related fashion, a network measured by reliability places a premium on ensuring that every critical element of the network has a low chance of failure (Siewert, 2005).

#### *e. Client Coverage Area*

Various techniques exist for quantifying and valuing client coverage area. For example, an area can be considered covered if it receives a signal strength greater than a certain threshold, if it supports a certain amount of client demand, or if it satisfies other criteria. The value of coverage can be determined by the worth assigned to a particular client or area, by surface area, or by other metrics.

Depending on the particular metric, greater coverage can be achieved by utilizing techniques similar to increasing network throughput: transmitting more power or at a higher frequency, or utilizing a larger bandwidth or a more complex modulation scheme. It can also be achieved by placing APs in elevated, central locations where radio emissions have a greater chance of being received, and by placing APs farther from each other to increase the total covered area.

Calegari et al., (2001), Raisanen and Whitaker (2005), Allen et al., (2002), Sharkey (2008) and He et al., (2004) all consider coverage in their respective research.

We adopt and modify the coverage shortfall concept of He et al., (2004) in this thesis. This technique is explained in detail in Chapter III.

## B. CURRENT WMN DESIGN TECHNIQUES

### 1. Traditional Solution: Trial-and-Error

The traditional method of designing WMNs involves trial-and-error site surveys. The basic idea is to place a small number of wireless APs in a portion of the desired coverage area, and then have a team walk or drive through the area, testing connectivity characteristics using network analyzers, client devices, or other tools as they move. When they achieve desired results for this sub-area, they add additional wireless APs to the network, and the process repeats (Olexa, 2005, pp. 146–147).

### 2. Commercial Solutions

There are several commercial high-fidelity tools for estimating radio propagation patterns for wireless APs. These include the USMC's System Planning, Engineering, and Evaluation Device (SPEED), and Motorola's MeshPlanner. The SPEED tool, developed by Northrop Grumman Mission Systems, receives as inputs Digital Terrain Elevation Data (DTED) maps, transmitter and antenna characteristics, modulation schemes, background radiation levels, and many other variables. Using the Terrain Integrated Rough Earth Model (TIREM) of signal propagation (Alion, 2007), SPEED can calculate point-to-point coverage characteristics between transmitters (such as signal to noise ratio (SNR), received signal strength, and theoretical bit error rate (BER)), as well as provide radio coverage analysis (RCA) around single wireless APs to determine client coverage areas.

Motorola's MeshPlanner has much of the same functionality as SPEED, and several major advantages. The program can utilize much more accurate environment information, including files in AutoCAD and shapefile formats. This allows the user to simulate the effects of buildings, vegetation and obstructions on the propagation pattern. The program also produces three-dimensional representations of wireless coverage,

assisting in designing networks to cover clients in elevated positions, such as towers or multistory buildings (Motorola, 2007).

In practice, network designers use these commercial tools and others in combination with site surveys and other tests. The results are then used to make adjustments to the input to the software tools, and the process repeats.

### 3. Shortfalls of Current Techniques

The trial-and-error technique has many drawbacks, the most critical of which to our intended purposes are the number of personnel and the amount of time required to obtain a working solution. These two resources are at a premium during both military and HA/DR operations.

Drawbacks to the SPEED program include a coarse 30-meter resolution of the terrain data—much too granular for most wireless AP coverage calculations – and the inability to include vegetation or manmade obstruction data.

Most notably, none of these techniques provides suggested wireless AP locations with any assurance of a good solution. A possible upgrade to the SPEED program by the Mitre Corporation may utilize simulated random walks to search for good connectivity solutions between APs (Shyy, 2008), and MeshPlanner generates AP locations automatically using heuristics (Vieira and Stoneback, 2008), but neither utilize an algorithm shown to converge to the global optimum. Essentially, engineers must design WMNs based on experience and trial-and-error using these limited support tools.

## C. THESIS OBJECTIVE: RATIONAL WWN DESIGN

We present a WMN design technique and associated tool that quickly provides good network topologies, using an algorithm with proven global convergence. The problem formulation (described in the next chapter) maximizes client coverage while considering many of the design constraints described here, including network performance, the quantity and technical characteristics of AP and client devices, and radio propagation physics. We further demonstrate these theoretical networks can be

practical in a real-world environment. The design tool can be operated with very little technical knowledge, and provides a useful step toward creating a rational, automated WMN design system.

Our formulation allows for a wide range of input parameters, but we purposefully avoid requiring assumptions about technology proprietary to a specific manufacturer or device. For instance, antennae of any height or gain may be considered, bandwidth can be of any quantity, and operating frequencies may range from 1 MHz to 20 GHz, but we do not consider the effects of modulation scheme or routing protocols. While this general focus perhaps reduces the predictive accuracy of our technique, it greatly increases its applicability. Nor is it impossible to modify our technique to incorporate such specific parameters, if available.

THIS PAGE INTENTIONALLY LEFT BLANK

# III. MODEL FORMULATION

## A. SOLUTION APPROACH

We seek the optimal locations of wireless mesh APs to maximize client coverage, subject to restrictions on network service, AP characteristics and placement, coverage requirements, and radio propagation over terrain. As noted earlier, there is an inherent tension between maximizing client coverage and network traffic flow. The crux of the SRRA+C formulation is to capture and quantify that tension.

We first describe each element of our formulation separately, building up to the entire SRRA+C problem statement. We then describe how to solve the SRRA+C problem utilizing the DIRECT algorithm. We also present an overview of the decision support tool we developed to facilitate solving the SRRA+C problem.

We define the *operating region* as the topographic area where the user may place APs. We discretize this operating region into a grid $(u,v) \in G$ where $u$ = 1,2,…$U$, $v$ = 1,2,…$V$. Each $(u,v) \in G$ ordered pair denotes a discrete area with a corresponding elevation in meters $E_{uv}$. We define the set $C$ (where $(u,v) \in C \subset G$) as the *coverage region*, that area within the operating region where client coverage is specifically desired. We define an indexed set $i \in N$, where $N$ is the set of all AP nodes {*1, 2,*…*n*}. We assume our network can be represented as a mesh of AP nodes *i* emanating from a single central node $d \in N$, also referred to as the gateway or headquarters node (see Figure 7 below). We assume the gateway node is the first node in the set *N*, and its position is fixed in advance. Once an AP has been placed, we assume it will not move.

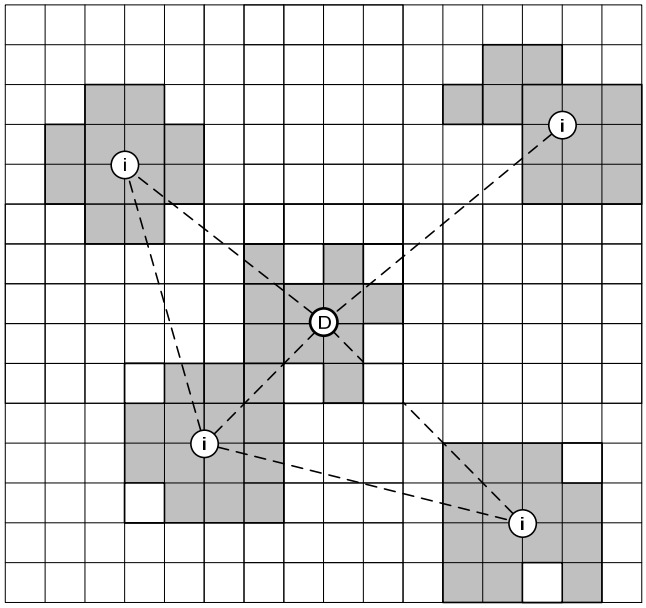


***Figure 7: Discretized operating area and WMN***

Each AP has two radios: one to provide client coverage (gray area), and the other to enable a backhaul network between nodes (dashed lines). Without loss of generality, we assume APs are not subject to interference from other APs.

We assume the vast majority of client network traffic will be directed to or from the gateway node *d*, as this location will connect to the Internet or other outside network, and will house email, domain, and storage servers. Traffic directly between AP nodes is permitted; we assume the nodes will route this traffic dynamically. Hence, we optimize our network for traffic flow from client service areas to the gateway node.

An alternate approach would be to optimize traffic flow from and to every node, to create peer-to-peer networks. While we don't specifically address this type of network, our formulation allows general traffic demands between any number of sources and/or destinations.

## B. CLIENT COVERAGE

### 1. Calculating Client Coverage

We assume the user will define the coverage region(s) $(u,v) \in C \subset G$ and the minimum signal strength required for a client to connect to an AP in that region. We also assume client devices and associated network demand are distributed uniformly within the defined coverage region. We desire to maximize total client coverage (which we define shortly) within the desired region. We assume client devices will connect to only one AP at any point in time, and each AP can support the client devices within its coverage region.

One basic method of calculating coverage utilizes "coverage circles." Using this technique, one would neglect the effects of terrain and other obstructions, and simply compute a coverage region for each AP based on the area of the circle circumscribed by a maximum coverage radius. This technique is computationally fast, but clearly represents an ideal (and very likely unrealistic) scenario, such as a network on a dry lake bed with perfect visibility in all directions. However, this method could be utilized to provide a "quick guess" towards a solution.

We utilize a more accurate method that calculates the *received signal strength* (RSS) at each discrete coverage location $(u,v) \in C$. Received signal strength $\rho$ at coverage region (*u*, *v*) from node *i* in dBm is calculated using the standard link budget formula (Olexa, 2005, p. 79):

$$\rho_{uv}^{i} = P_{tx} + g_{tx} - L_{tx} - L_{fs} - L_{m} + g_{rx} - L_{rx} \tag{1}$$

where $P_{tx}$ is transmitted power in dBm, $g_{tx}$ and $g_{rx}$ are respectively the gains of the AP and receiver in dBi, $L_{tx}$ and $L_{rx}$ are respectively the losses (i.e., from cables, connectors, etc.) of the AP and receiver in dB, $L_{fs}$ is free space loss in dB, and $L_m$ is miscellaneous loss (such as fade margin) in dB. We define transmission power between any two APs as

a decision variable, and gains, equipment losses, and fade margins as data. As all terms are in decibel form and the $P_{tx}$ term is a logarithmic function of power in watts, this function is monotonic increasing in $P_{tx}$.

Our formulation allows any method of computing free space loss $L_{fs}$. In this thesis, we consider three common methods used in network design. The first method computes loss using an inverse-square law. The loss is simply the inverse of the distance squared. While this method captures some aspect of the diffusion of radio wave power over distance, it is in no way an accurate method of determining loss and is used simply for testing purposes.

The second method of determining free space loss is through the Hata COST–231 model for radio propagation (European Cooperation in the Field of Scientific and Technical Research, 1999):

$$L = 46.3 + 33.9 \log f - 13.82 \log h_S - C_H + (44.9 - 6.55 \log h_S) \log 1000d + C$$

where $L$ is median path loss in dB, $f$ is transmission frequency in megahertz, $h_S$ is sender effective height in meters, $d$ is distance in meters, $C$ is a constant in dB (0 for medium cities and suburban areas, 3 for metropolitan areas), and $C_H$ is the receiver antenna height correction factor:

$$C_H = 0.8 + (1.1 \log f - 0.7) h_R - 1.56 \log f \qquad \text{for small \& medium cities}$$

$$C_H = \begin{cases} 8.29(\log(1.54 h_R))^2 - 1.1, & \text{if } 150 \le f \le 200 \\ 3.2(\log(11.75 h_R))^2 - 4.97, & \text{if } 200 \le f \le 1500 \end{cases} \qquad \text{for large cities}$$

where $h_R$ is receiver antenna height in meters. This method provides a good approximation of propagation loss and is computationally very quick. However, it does not specifically account for path obstructions, i.e., it does not consider terrain, and it is valid only for frequencies between 1.5 and 2.0 GHz.

A third method utilizes the Terrain Integrated Rough Earth Model (TIREM) of Alion Science & Technology Corporation (Alion, 2007). This commercial tool is the same propagation method used in the Marine Corps' SPEED tool and other commercial planning tools such as the Satellite Toolkit of Analytical Graphics, Inc. (2009). The

model utilizes terrain information to compute a path profile between sender and receiver, sampling terrain elevation at a given frequency. This path profile is then used to compute propagation loss, considering the effects of free space loss, diffraction around obstacles, and atmospheric absorption and reflection. TIREM also considers the following variables:

- Transmission frequency
- Transmitter and receiver antenna gain, height, and polarization
- Conductivity of the earth's surface
- Surface humidity
- Relative permittivity of the earth's surface
- Surface refractivity

The model has a wide range of valid input values; for instance, frequency can vary between 1 MHz and 20 GHz, and antenna height can vary between ground level and 30 kilometers. While this method is the most accurate of the three, it is also the most computationally expensive.

A model we do not utilize is the Irregular Terrain Model (Longley and Rice, 1968). This model is based on both statistical analyses and actual measurements, and is accurate for frequency ranges from 20 MHz to 20 GHz (National Telecommunications and Information Administration, 2008). The model is freely available and is very popular in this area of research, but is not valid for distances of less than 1 kilometer and hence not suited for most WMN applications.

### 2. Calculating the Value of Client Coverage

In principle, one can compute the value of client coverage as a simple sum of the number of discrete regions, which receive a minimal threshold of service. In this fashion, each region would or would not be covered, indicated by a binary decision variable. This method is simple to understand, but the binary decision variables greatly increase computational complexity.

To avoid creating a mixed integer program, we adopt and modify He et al.'s (2004) concept of power coverage. We assume each coverage region (which need not be contiguous) has a minimum *coverage threshold* $\tau_{uv}$ in watts. Any received signal above this threshold results in useable client coverage. For each $(u,v) \in C$, we calculate $\rho_{uv}{}^{i}$, the actual received signal strength from node *i* using Equation (1). The difference of $\rho_{uv}{}^{i}$ (converted to watts) and $\tau_{uv}$ represents a quantity defined as *coverage shortfall* at region (*u, v*) from node *i*. We wish to minimize this quantity: a positive difference represents inadequate power coverage. Additionally, we can indicate regions where greater network demand is expected (or a region of greater importance) by utilizing larger threshold values $\tau_{uv}$. Summarizing, we have

$$\left(\text{Coverage Shortfall}\right)_{uv}^{i} \equiv \left(\tau_{uv} - \rho_{uv}^{i}\right).$$

As each client device can connect to only one AP at any point in time, we consider only the minimum coverage shortfall calculated from each node *i*. We allow only positive terms, to remove any benefit of "blasting" a particular region with power. This yields:

$$\left(\text{Coverage Shortfall}\right)_{uv} \equiv \min\left(\tau_{uv} - \rho_{uv}^{i}\right)_{+}$$

where $()_{+}$ denotes the projection onto the nonnegative orthant. We desire to find the coverage shortfall for all possible coverage grid squares $(u,v) \in C$, so we sum over all $(u,v) \in C$. This yields total coverage shortfall:

$$\left(\text{Total Coverage Shortfall}\right) \equiv \sum_{(u,v)\in C} \min\left(\tau_{uv} - \rho_{uv}^{i}\right)_{+}. \qquad (2)$$

This measure avoids double counting when a particular $(u,v) \in C$ receives more than $\tau_{uv}$ from more than one AP. It also bounds the amount any region can be penalized: no more than $\tau_{uv}$.

We use total coverage shortfall in our formulation. Other coverage objectives could involve average coverage shortfall (obtained by dividing by the cardinality of the set of all coverage grid squares *C*), and maximum coverage shortfall:

$$\left(\text{Average Coverage Shortfall}\right) \equiv \frac{1}{|C|}\sum_{(u,v)\in C} \min\left(\tau_{uv} - \rho_{uv}^{i}\right)_{+}$$

$$\left(\text{Maximum Coverage Shortfall}\right) \equiv \max\left[\min\left(\tau_{uv} - \rho_{uv}^{i}\right)_{+}\right].$$

## C. NETWORK ROUTING AND RESOURCE ALLOCATION

We formulate a network flow problem to maximize the delivered traffic in bits per second (bps) from any node *i* to the HQ node, given constraints on communications resources. First, we calculate arc capacities between each node and then use this data in our network problem.

### 1. Calculating Arc Capacities

In addition to calculating client coverage, we also need to calculate the flow of traffic between AP nodes as a measure of network performance. We use a network flow problem (Chapter III.C.3) to optimize these flows, maximizing the flow in the network from any node *i* to the gateway (HQ) node.

Flow maximization requires as input arc capacities between AP nodes. In computing arc capacity, we can ignore many of the effects of a particular device's actual hardware and software characteristics by using the Shannon capacity formula (1949), which establishes a theoretical upper bound on radio link capacity in bits per second:

$$(Capacity)_{ij} = B\log_2\left(1+\frac{Signal_{ij}}{b_{ij}}\right) \quad \forall (i,j)\in A \tag{3}$$

where $B$ is the channel bandwidth in hertz, and $Signal_{ij}$ and $b_{ij}$ are respectively the signal power and background noise power in watts or volt$^2$ from node *i* to node *j*. We desire to calculate the received signal capacity, so the *Signal* term is calculated using the antilog of the link budget formula (Equation 1). Following Xiao et al., (2004), this yields:

$$(Capacity)_{ij} = B\log_2\left(1+\frac{g_{ij}}{b_{ij}l_{ij}}P_{ij}\right) \quad \forall(i,j)\in A \tag{4}$$

where $g_{ij}$ is the sum of the antilog gain terms, $l_{ij}$ is the sum of the antilog loss terms, and $P_{ij}$ represents the fraction of a finite resource (in this case, transmission power in watts) used to transmit from *i* to *j*. We assume

$$\sum_{j:(i,j)\in A} P_{ij} \le p \tag{5}$$

that is, the sum of all resource fractions is less than or equal to a constant representing the total available resource.

### 2. Calculating the Value of Delivered Network Flow

We measure an individual traffic flow in bits per second. However, we need to assess the value of all flows across the network. One approach is to simply compute the sum of all delivered traffic from all APs to the HQ node. Let $S_i^d$ be the total flow of traffic from source node $i \in N$ to destination node $d \in D$. Then

$$\left(Total\ Network\ Flow\right) \equiv \sum_d \sum_{i\neq d}\left(S_i^d\right).$$

In this case, all traffic is valued equally, independent of other traffic levels. As long as the total volume of traffic is large; it doesn't matter which source-destination pairs are passing traffic (or not).

Another approach is to consider minimizing network flow shortfall, in a fashion similar to our coverage shortfall formulation. Let *min_flow* be the minimum flow delivered between any outlying AP and the HQ node. Then

$$\left(Total\ Network\ Flow\ Shortfall\right) \equiv \sum_d \sum_{i\neq d}\left(min\_flow - S_i^d\right)_+ .$$

This objective function is appealing because desired network performance can be specified through a concrete number. However, our testing reveals that this method causes problems when solving our network formulation (described in the next section). Specifically, the ability of our problem to converge is extremely sensitive to choices of *min_flow*. For these reasons, we abandon consideration of this alternative method.

An alternate approach adopted by Xiao et al., (2004) is to consider the utility of delivered traffic. Specifically, they use a log-utility function which places a zero value on unit flow, positive values on flows greater than one, and negative values on flows less than one:

$$\left(\textit{Utility of Total Network Flow}\right) \equiv \sum_{d} \sum_{i \neq d} \log_2 \left( S_i^d \right).$$

This utility function promotes *fairness* between flows: the penalty for small flows encourages the assignment of at least unit flow to all source-destination pairs.

### 3. Simultaneous Routing & Resource Allocation Formulation

We use Xiao et al.,'s (2004) Simultaneous Routing and Resource Allocation (SRRA) technique to calculate a value of network flow. Our goal is to maximize the utility of network flow from each node to the central HQ node. We assume AP locations are given, the communications terms $g_{ij}$ and $l_{ij}$ are pre-calculated, and total transmission resource $p$ is a fixed quantity representing transmission power in watts available at each AP.

**SRRA Formulation**

Index Use

| | |
|---|---|
| $i \in N$ | node *(alias j, k)* |
| $(i, j) \in A$ | arc *(link)* |
| $d \in D \subseteq N$ | destination node |

Calculated Data

| | | |
|---|---|---|
| $g_{ij}$ | product of antilog gain terms from $i \in N$ to $j \in N$ | [none] |
| $l_{ij}$ | product of antilog loss terms from $i \in N$ to $j \in N$ | [none] |
| $p_i$ | maximum total transmission power per node | [watts] |
| $B$ | channel bandwidth | [hertz] |
| $b_{ij}$ | background noise power from $i \in N$ to $j \in N$ | [watts] |

Decision Variables

| | | |
|---|---|---|
| $S_i^d$ | total flow of traffic from $i \in N$ to destination $d \in D$ | [flow] |
| $X_{ij}^d$ | traffic flow along arc $(i, j) \in A$ to destination $d \in D$ | [flow] |
| $T_{ij}$ | total flow vector along arc $(i, j) \in A$ | [flow] |
| $P_{ij}$ | total transmission power along arc $(i, j) \in A$ | [watts] |

Formulation

$$\max_{S,X,T,P} \sum_d \sum_{i \neq d} \log_2 \left( S_i^d \right) \qquad \text{(S0)}$$

s.t.

$$\sum_{k:(j,k) \in A} X_{jk}^d - \sum_{i:(i,j) \in A} X_{ij}^d = S_j^d \qquad \forall j \in N, \forall d \in D \qquad \text{(S1)}$$

$$T_{ij} = \sum_d X_{ij}^d \qquad \forall (i,j) \in A \qquad \text{(S2)}$$

$$T_{ij} - B \log_2 \left( 1 + \frac{g_{ij}}{b_{ij} l_{ij}} P_{ij} \right) \leq 0 \qquad \forall (i,j) \in A \qquad \text{(S3)}$$

$$\sum_{j:(i,j) \in A} P_{ij} \leq p_i \qquad \forall i \in N \qquad \text{(S4)}$$

$$P_{ij} \geq 0 \qquad \forall (i, j) \in A \qquad \text{(S5)}$$

$$X_{ij}^d \geq 0 \qquad \forall (i, j) \in A, \ \forall d \in D \qquad \text{(S6)}$$

$$T_{ij} \geq 0 \qquad \forall (i, j) \in A \qquad \text{(S7)}$$

$$S_i^d \geq 0 \qquad i \neq d \qquad \text{(S8)}$$

Discussion

The objective function (S0) maximizes the sum of utilities for each traffic flow. Constraint (S1) ensures balance of flow at each node. Constraint (S2) defines the total flow along any arc to equal the sum of all traffic flows along that arc. Constraint (S3)

ensures traffic flow along any arc is less than or equal to the capacity of that arc. Constraint (S4) bounds total transmission power per AP by available resource. Constraints (S5-S8) ensure non-negativity.

## 4. Dual Decomposition Technique

Xiao et al., (2004) observe that due to its layered structure (i.e., the routing and resource allocation decisions), this problem can be solved via dual decomposition. We follow their construction. By introducing the Lagrange multipliers $\alpha_{ij} \forall (i,j) \in \text{A}$ for the routing and resource allocation constraint (S3), we obtain the partial Lagrangian

$$L(S,X,T,P,\alpha) = \sum_{d}\sum_{i \neq d} \log_2(S_i^d) - \sum_{(i,j)\in A} \alpha_{ij}\left(T_{ij} - B\log_2\left(1+\frac{g_{ij}}{b_{ij}l_{ij}}P_{ij}\right)\right). \quad (6)$$

We minimize this by choice of α:

$$\min_{\alpha} L(S,X,T,P,\alpha).$$

The objective function of the dual problem can then be evaluated via the network flow variables *S*, *X*, and *T*, and the communications variable *P*:

$$V(\alpha) = V_{flow}(\alpha) + V_{comm}(\alpha) \quad (7)$$

$$V_{flow}(\alpha) = \max_{S,X,T} \sum_{d}\sum_{i \neq d} \log_2(S_i^d) - \sum_{(i,j)\in A} \alpha_{ij} T_{ij} \quad (8)$$

$$V_{comm}(\alpha) = \max_{P} \sum_{(i,j)\in A} \alpha_{ij} B\log_2\left(1+\frac{g_{ij}}{b_{ij}l_{ij}}P_{ij}\right) \quad (9)$$

The representation of the capacity constraint (Equation 4) is a logarithmic function and hence meets the assumptions of Xiao et al., (2004) of a concave and monotone increasing capacity function. Xiao et al., (2004) also assume that Slater's condition for constraint qualification is satisfied:

$$T_{ij} < B \log_2 \left( 1 + \frac{g_{ij}}{b_{ij} l_{ij}} P_{ij} \right) \quad \forall (i, j) \in A \tag{10}$$

that is, the total flow of traffic along a link is strictly less than the maximum capacity of that link. They claim this assumption is "almost always true in practice." Combined with their observation that the dual function $V(\alpha)$ (Equation 7) is always convex, Xiao et al., (2004) conclude that strong duality holds, i.e., the solution to the dual and primal problems are equal.

Xiao et al., (2004) observe that since the objective function of the primal problem is not strictly concave in the variables $X$ and $T$, the dual function is only piecewise differentiable, and hence a non-differentiable convex optimization problem. They apply the subgradient method to obtain a solution. Similar to the gradient method, this method uses the subgradient rather than the gradient (which may not exist or is too difficult to compute). Each iteration of the subgradient method might not necessarily improve the dual objective value, but each iteration reduces the distance to the optimal solution (Bertsekas, 1999, p. 621).

Xiao et al., (2004) evaluate each component of $V(\alpha)$ separately. The solution of $V_{flow}(\alpha)$ (Equation 8) is computed via the *network flow subproblem*, and the solution of $V_{comm}(\alpha)$ (Equation 9) is computed via the *resource allocation subproblem*. Xiao et al., (2004) observe that both subproblems are convex optimization problems with special structure lending themselves to very efficient computational techniques.

### *a. Network Flow Subproblem*

We solve the network flow subproblem (defined below) as a multicommodity flow problem.

**Network Flow Subproblem**

$$\max_{S,X,T} \sum_{d} \sum_{i \neq d} \log_2(S_i^d) - \sum_{(i,j) \in A} \alpha_{ij} T_{ij} \qquad \text{(F0)}$$

$$s.t. \quad \sum_{k:(j,k) \in A} X_{jk}^d - \sum_{i:(i,j) \in A} X_{ij}^d = S_j^d \qquad \forall j \in N, \forall d \in D \qquad \text{(F1)}$$

$$T_{ij} = \sum_{d} X_{ij}^d \qquad \forall (i,j) \in A \qquad \text{(F2)}$$

$$X_{ij}^d \geq 0 \qquad \forall (i,j) \in A, \ \forall d \in D \qquad \text{(F3)}$$

$$T_{ij} \geq 0 \qquad \forall (i,j) \in A \qquad \text{(F4)}$$

$$S_i^d \geq 0 \qquad i \neq d \qquad \text{(F5)}$$

Discussion

The objective function (F0) maximizes the difference of the sum of utilities for each traffic flow and the costs associated with the total flows along each arc. Constraint (F1) ensures balance of flow at each node. Constraint (F2) defines the total flow along any arc to equal the sum of all traffic flows along that arc. Constraints (F3-F5) ensure non-negativity.

By substituting the $T_{ij}$ term in the objective function with $\sum_d X_{ij}^d$, the problem can be solved for each fixed commodity (destination) $\hat{d}$:

$$\max_{S,X,T} \sum_{i \neq \hat{d}} \log_2(S_i^{\hat{d}}) - \sum_{(i,j) \in A} \alpha_{ij} X_{ij}^{\hat{d}}$$

$$s.t. \quad \sum_{k:(j,k) \in A} X_{jk}^{\hat{d}} - \sum_{i:(i,j) \in A} X_{ij}^{\hat{d}} = S_j^{\hat{d}} \qquad \forall j \in N$$

$$X_{ij}^{\hat{d}} \geq 0 \qquad \forall (i,j) \in A$$

$$S_i^{\hat{d}} \geq 0 \qquad i \neq \hat{d}$$

We next consider the total cost of sending flow over different paths from source node *i* to destination node *d*. Let $B_k$ denote the $k^{th}$ path from *i* to *d*, and let $Y_k$ denote the total flow sent along path $B_k$. Then

$$(\textit{Total Cost Per Unit Flow})_{B_k} = \sum_{(i,j) \in B_k} \alpha_{ij}$$

and

$$(Total\ Cost\ of\ Flow)_{B_k} = Y_k \sum_{(i,j)\in B_k} \alpha_{ij} .$$

Summing over all paths from $i$ to $d$, the contribution to the objective function can be rewritten as

$$\log_2\left(\sum_k Y_k\right) - \sum_k Y_k \sum_{(i,j)\in B_k} \alpha_{ij} .$$

This objective is maximized when the marginal utility equals the marginal cost. For any path $B_k$, this maximum is reached at the point

$$\frac{d}{dY_k}\log_2 Y_k = \sum_{(i,j)\in B_k} \alpha_{ij} .$$

*Proposition*: For each source-destination pair ($i$,$d$), the optimum solution is to send flow only along the shortest (lowest cost) path(s) from $i$ to $d$. Thus, solving for optimal flows from $i$ to $d$ is achieved by calculating the shortest path(s) and solving for first order conditions.

*Proof*: Consider two paths $B_1$ and $B_2$ from $i$ to $d$ with respective costs $c_1$ and $c_2$, where $c_1 < c_2$. Flow $Y_1$ should be pushed along path $B_1$ as long as the marginal benefit is greater than the marginal cost, i.e.,

$$\frac{d}{dY_1}\log_2 Y_1 \geq c_1 .$$

The stopping point is reached at equality, as pushing additional flow beyond this point results in greater cost than benefit. Increasing delivered flow along $B_2$ costs at least $c_1$ per unit (since $c_2 > c_1$), and generates marginal benefit no greater than $c_1$. Thus, the optimal solution is to push $Y_1^*$ only along path $B_1$.

If multiple shortest paths exist (e.g., there exists a path $B_3$ for which $c_3 = c_1$), the optimal solution is to push any convex combination of nonnegative flows totaling $Y_1^*$ along these shortest paths, and no flows along strictly longer paths. Q.E.D.

The ability to send flow along multiple shortest paths (if they exist) is required for our problem to converge under the subgradient method. We use a modification of the Floyd-Warshall algorithm that can store all shortest paths from *i* to *d*, and use the associated path costs to solve for the optimal flows

$$Y_k^* = S_i^d = \frac{1}{c_k \ln(2)}$$

where $c_k$ is the lowest cost path *k*, namely, the smallest α between node *i* and *d*. If multiple shortest paths do exist for any (*i, d*), we split the optimal flow evenly among the paths.

The following pseudo-code solves the network flow subproblem.

**Algorithm Network Flow**

Input: Node and arc adjacency lists; maximum transmit power *p*; α values for each arc.

Output: Network flow values for all transmission links

**begin**
    Calculate arc costs
    Calculate all shortest paths for all $(i, j) \in A$ using Floyd-Warshall algorithm
    **for** (all (*i, d*))
        Calculate total delivered flow $S_i^d$ from each $i \in N$ to all $d \in D$
        Calculate flow $X_{ij}^d$ along each $(i, j) \in A$ to all $d \in D$

    **next** (*i, d*)
**end;**

### *b. Resource Allocation Subproblem*

The resource allocation subproblem determines transmitted power $P_{ij}$ and can be expressed as follows:

**Resource Allocation Subproblem**

$$\max_{P} \sum_{(i,j)\in A} \alpha_{ij} B \log_2\left(1+\frac{g_{ij}}{b_{ij}l_{ij}}P_{ij}\right) \quad \text{(R0)}$$

$$s.t. \sum_{j:(i,j)\in A} P_{ij} \le p_i \qquad \forall i \in N \quad \text{(R1)}$$

$$P_{ij} \ge 0 \qquad \forall (i,j) \in A \quad \text{(R2)}$$

Discussion

The objective function (R0) maximizes the sum of the capacities of each arc. Constraint (R1) bounds total transmission power per AP by available resource. Constraint (R2) ensure non-negativity.

We solve this nonlinear, concave, differentiable problem by using a resource allocation algorithm described by Luss and Gupta (1974). The method allows us to solve for each AP power allocation scheme separately. Let $k$ = 1, 2,…$R$ be the set of all adjacent arcs $(i, j) \in A$ to the AP in question. Let $a_k$ and $c_k$ represent the lower and upper bounds of $P_k$. In the above formulation, $a_k$ = 0 and $c_k = p_i$. Let $DQ(P_k)$ be the derivative of the objective function at $P_k$.

By applying the Kuhn-Tucker conditions, Luss and Gupta (1974) derive that a necessary and sufficient condition for $P^*$ to be an optimal solution is that there exists a Lagrange multiplier $M$ such that:

$$P_k^* = a_k \rightarrow DQ(a_k) \le M \quad (11)$$

$$a_k < P_k^* < c_k \rightarrow DQ(P_k^*) = M \quad (12)$$

$$P_k^* = b_k \rightarrow DQ(c_k) \ge M \quad (13)$$

$$\sum_{k=1}^{R} P_k = p_i \quad \forall i \in N. \quad (14)$$

Equations (11-13) establish bounds on $M$ based on the optimum value of $P$. Equation (14) requires the sum of all transmission powers to equal the total possible transmission power, i.e., all available transmission resources are utilized.

Luss and Gupta (1974) show that if $DQ(a_k) \geq DQ(a_{k+1})$ for $k$ = 1, 2,…$R$-1, then there exists an integer $q$, $1 \leq q \leq R$, such that $P_k^* > a_k$ , $k$=1, 2,…$q$, and $P_k^* = a_k$ , $k$=$q$+1, $q$+2,…$R$. By reordering the derivatives $DQ(a_k)$ in this manner, we ensure $P_k^* > a_k$ (the optimal power allocation is strictly greater than the lower bound of zero) and $DQ(P_k^*) = M(q)$, where

$$M(q) = (1/q)\sum_{k=1}^{R} DQ\left(P_k^*\right).$$

The optimal solution $P_k^*$ can therefore be found by finding a closed-form expression for $M(q)$. First, we solve for $P_k^*$ as a function of $M(q)$. The derivative of the objective function is

$$DQ(P_k) = \frac{\alpha_k B \dfrac{g_k}{b_k l_k}}{\ln(2)\left(1 + \dfrac{g_k}{b_k l_k} P_k\right)} \quad \forall k \in R .$$

Since the derivatives are ordered, $DQ(P_k^*) = M(q)$ and we solve for $P_k^*$ as a function of $M(q)$:

$$P_k^* = \frac{\alpha_k B}{\ln(2) M(q)} - \frac{b_k l_k}{g_k} \quad \forall k \in R .$$

To find a closed-form expression for $M(q)$, we first simplify the derivative:

$$M(k) = DQ(P_k^*) = \frac{\alpha_k B}{\ln(2)\left[P_k^* + \left(\dfrac{b_k l_k}{g_k}\right)\right]} \quad \forall k \in R .$$

Next, we sum over all $k$=1, 2,…$q$:

$$M(q) = \frac{B\sum_{k=1}^{q} \alpha_k}{\ln(2)\sum_{k=1}^{q} P_k^* + \ln(2)\sum_{k=1}^{q}\left(\frac{b_k l_k}{g_k}\right)}.$$

By the above results, we can replace $\sum_{k=1}^{q} P_k^*$ with $p_i - \sum_{k=q}^{R} a_k$, yielding

$$M(q) = \frac{B\sum_{k=1}^{q} \alpha_k}{\ln(2)\left[p_i - \sum_{k=q}^{R} a_k\right] + \ln(2)\sum_{k=1}^{q}\left(\frac{b_k l_k}{g_k}\right)}.$$

Since the lower bound on transmission power $a_k = 0$, this expression simplifies to

$$M(q) = \frac{B\sum_{k=1}^{q} \alpha_k}{\ln(2)\left[p_i + \sum_{k=1}^{q}\left(\frac{b_k l_k}{g_k}\right)\right]}.$$

To solve for the optimum value of $q$, one must iteratively increase $q$ by one until the condition

$$DQ\left(a_{k+1}\right) > M(q)$$

is no longer true, or until $k$ equals the number of adjacent arcs $R$.

Building on Luss and Gupta (1974), the following pseudo-code solves the resource allocation subproblem.

**Algorithm Resource Allocation**

Input: Node and arc adjacency lists; maximum transmit power *p*; gain, loss and α values for each arc; channel bandwidth *B*, background noise *b*

Output: Power allocations for all arcs

**begin**
Reorder derivatives
Set k = 0
**do**
k = k + 1
Calculate $M(q)$
**if** ($k < R$)
Calculate $DQ(a_{k+1})$
**end if;**
**while** ( $DQ(a_{k+1}) > M(q)$ and $k < R$)
**end;**
Calculate $P_k^*$ for all adjacent arcs
**end;**

### *c. Utilizing the Subgradient Method*

The subgradient of the non-differentiable convex function $V(\alpha)$ is a vector *h* such that

$$V(\beta) \geq V(\alpha) + h'(\beta - \alpha) \quad \forall \beta .$$

Let $X^*(\alpha)$, $T^*(\alpha)$, and $S^*(\alpha)$ be an optimal solution to the network flow subproblem at $\alpha$, and let $P^*(\alpha)$ be an optimal solution to the resource allocation subproblem at $\alpha$. Then the subgradient *h* of $V(\alpha)$ is

$$h_{ij} = B \log_2 \left( 1 + \frac{g_{ij}}{b_{ij} l_{ij}} P_{ij}^*(\alpha) \right) - T_{ij}^*(\alpha)$$

where $h_{ij}$ are the components of the subgradient *h* defined for each $(i, j) \in A$. One can interpret the subgradient as the excess capacity on each arc.

To solve the dual problem via the subgradient method, we select an initial value of $\alpha^1$ (we use $\frac{g_{ij}}{b_{ij} l_{ij}}$), and for each iteration $k$ = 1, 2, 3, …, we compute the dual function $V(\alpha^k)$ and subgradient $h^k$. We then update the dual variable $\alpha$ according to

$$\alpha^{k+1} = \left(\alpha^k - \delta_k h^k\right)_+$$

where $\delta_k$ is the stepsize of iteration $k$ and $()_+$ denotes the projection onto the nonnegative orthant. We use a stepsize rule of

$$\delta_k = \frac{a^k}{|| h_{ij} ||^2}$$

where

$$a^k = \frac{1+m}{k+m}$$

and $m$ is a fixed positive integer (we use 1000). This stepsize rule follows the conditions for a diminishing stepsize (Bertsekas, 1999, p. 624), namely

$$a^k \to 0, \quad \sum_{k=1}^{\infty} a^k = \infty .$$

We stop the method after a given number of iterations (typically 500), and use the best solution found as an approximation of the solution to the SRRA problem (and an estimate of network performance). We explain this in further detail in Chapter III.D.

### 5. SRRA Algorithm

Building on Xiao et al., (2004), the following pseudo-code solves the simultaneous routing and resource allocation problem using the subgradient method.

**Algorithm SRRA**

Input: Stopping criteria (*max_iterations*)

Output: Optimal routing and power allocations for all transmission links

**begin**

Select initial $\alpha$ values

**while** ($k <$ *max_iterations*)

**call *Algorithm Network Flow*** (calculates $X_{ij}^{d}$ values)

Calculate $T_{ij} = \sum_{d} X_{ij}^{d} \quad \forall (i,j) \in A$

Calculate $S_i^d \quad \forall i \in N, \forall d \in D$

Calculate $V_{flow}(\alpha) = \sum_{d} \sum_{i \neq d} \log_2 (S_i^d) - \sum_{(i,j) \in A} \alpha_{ij} T_{ij}$

**call *Algorithm Resource Allocation*** (calculates $P_{ij}$ values)

Calculate $V_{comm}(\alpha) = \sum_{(i,j) \in A} \alpha_{ij} B \log_2 \left( 1 + \frac{g_{ij}}{b_{ij} l_{ij}} P_{ij} \right)$

Calculate $h_{ij} = B \log_2 \left( 1 + \frac{g_{ij}}{b_{ij} l_{ij}} P_{ij} \right) - T_{ij}^{*}(\alpha) \quad \forall (i,j) \in A$

Calculate $\delta_k$

Calculate $\alpha^{k+1} = \left( \alpha^k - \delta_k h^k \right)_+$

**end;**

**end;**

## D. SRRA+C FORMULATION

### 1. Weighted Multiple Objective Function

We return to our overall objective of maximizing coverage area (or more specifically, minimizing coverage shortfall), subject to constraints on number and operating characteristics of APs, network flow requirements, and the operating environment. The DIRECT optimization algorithm (described below) requires the objective function to be continuous; to meet this requirement, we enforce a lower bound on network service using a soft constraint. Greater network flow results in greater benefit (i.e., smaller penalty) to the objective function:

$$\left( Overall\ Objective\ Function \right) \equiv \left( Total\ Coverage\ Shortfall \right) - w\left( Network\ Flow \right)$$

where $w$ is a positive scalar (in units of dBm/log$_2$bps) representing the relative weight placed on network flow. Larger values of $w$ place more emphasis on network flow, and in general increase the desirability of more compact network topologies. Clearly, the order of magnitude of $w$ depends on the relative values of coverage shortfall and network flow.

We solve total coverage shortfall using the definition in Chapter III.B, and network flow using the SRRA formulation. Both are functions of AP locations. Let $\sigma$ be the locations of all APs for a particular network, and let $\Omega$ be the overall objective function. Thus,

$$\Omega(\sigma) \equiv \sum_{(u,v)\in G} \min(\tau_{uv} - \rho^{i}_{uv})_{+} - w \max_{S,X,T,R} \sum_{d} \sum_{i \neq d} \log_2(S_i^d). \qquad (15)$$

The SRRA+C problem is to minimize the objective value $\Omega$ by choice of AP locations $\sigma$. The objective function is in units of dBM, although the combined objective value has no direct practical implication. Rather, the objective value serves as a relative method of comparing different network topologies.

Note that this formulation does not directly relate the values of client coverage and backhaul network flow. This follows from our assumption that all APs have different radio systems to provide client coverage and backhaul connectivity. These radios have separate resource constraints and can utilize different networking algorithms. This implies that a change to client coverage area (even if client distribution is uniform) does not necessarily result in a change to backhaul network flow: APs can have proprietary algorithms that may reduce client coverage area with demand, or maintain the area by reducing the amount of service each client receives.

### 2. Bounding the SRRA+C Problem

Four values bound the overall objective function in Equation (15), namely the respective upper and lower bounds of coverage shortfall and network flow. These help us to understand the limits to the possible improvement of any solution, (though as we show, this boundary may be infeasible).

At one extreme is "perfect" client coverage: every discrete grid square $(u,v) \in C$ receives an RSS of at least $\tau_{uv}$, so the coverage shortfall value is zero. At the other extreme, we have the exact opposite: no grid square receives any signal, so the coverage shortfall value is $\sum_{(u,v)\in C} \tau_{uv}$.

The lower bound on total delivered network flow $\sum_{d} \sum_{i \neq d} S_i^d$ is zero. The upper bound depends on the capacity of the given network with virtually no propagation loss between APs. (This is equivalent to placing all APs at essentially the same location.)

These four bounds provide a two-dimensional representation of the solution space (see Figure 8). The horizontal axis represents the value of coverage, and the vertical axis represents the value of network flow. A solution on the left vertical axis has perfect coverage, and a solution on the upper horizontal axis has perfect network flow. Note that it is highly unlikely for a solution to appear at the upper-left corner of this bounded solution space. Such a "perfect" solution may consist of a very small coverage region receiving service from an overabundance of co-located APs, or perhaps an arbitrary coverage region literally completely covered in APs. We consider this theoretical lower bound in Chapter IV.B.4, and in Chapter IV.B.5 we create a Pareto frontier in this solution space that can provide useful information on the value of any solution.

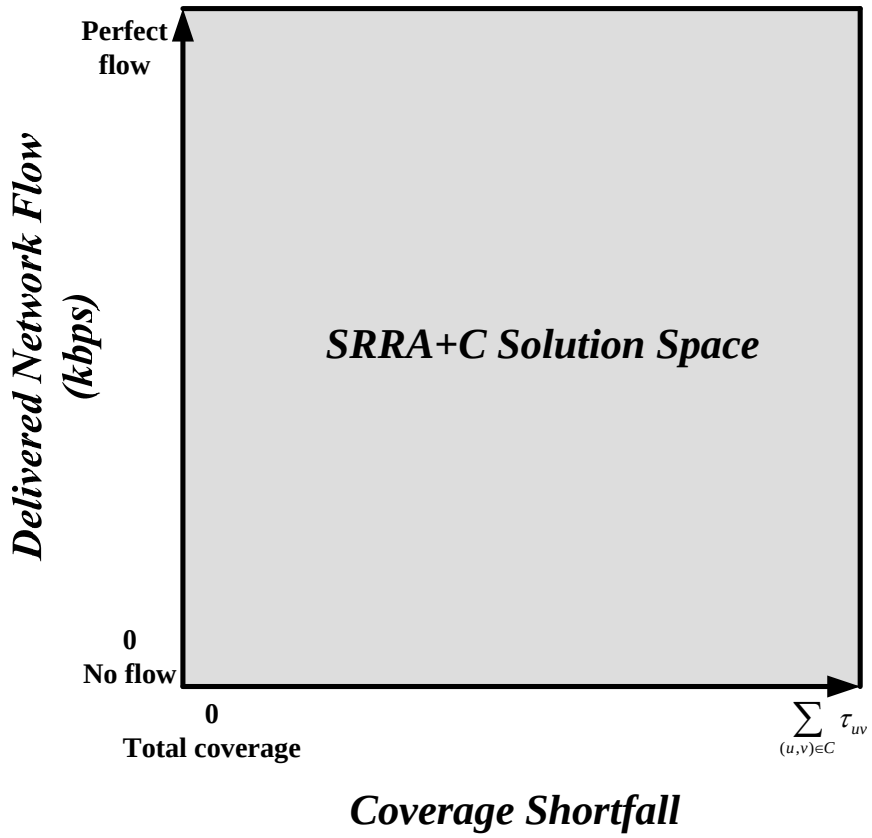


***Figure 8: The boundaries of the SRRA+C solution space***

### 3. Solving the SRRA+C Problem

Thus far, we have described how to quantify the value of a network using the SRRA+C formulation given AP locations, but not how to actually choose those locations. For a finite number of discrete AP locations, an exact method of solving the SRRA+C problem is through enumeration, i.e., trying all possible AP placement solutions. We show this technique is very inefficient.

Recall that the operating area (that area where APs may be placed and coverage may be desired) is discretized into a $U$ x $V$ grid. There are $n$ total APs and the position of the first (the HQ node) is given, so the total number of unique AP locations $\sigma$ is $\binom{UV}{n-1}$. This exponential growth in both $UV$ and $n$ of the solution space makes for a very large

number of solutions. For the purposes of illustration we consider a small operating region of 20x20 = 400 grid squares. The below table shows the growth of the solution space with the number of APs, and the time required to solve the associated SRRA+C problem using our decision support tool (described in Chapter III.G) via enumeration using a Windows XP laptop with an Intel Core 2 Duo processor operating at 2.16 GHz and 2 GB of RAM:

| APs | Number of Solutions | Processing Time |
|---|---|---|
| 2 | 400 | 3 sec, 415ms |
| 3 | 79,800 | 17 min, 1 sec, 504 ms |
| 4 | 10,586,800 | 2 days, 11 hours, 9 min, 53 sec, 523 ms |
| 5 | 1,050,739,900 | 245 days, 5 min *(extrapolated)* |

***Table 1: Exponential growth of processing time using enumeration***

Clearly we need a faster, more efficient solution method if SRRA+C is to be incorporated into the network design process. The DIRECT algorithm meets this need.

## E. DIVIDING RECTANGLES OPTIMIZATION

The DIviding RECTangles (DIRECT) algorithm of Jones et al., (1993) is a sampling optimization algorithm. The approach is similar to Lipschitzian optimization (Horst and Hoang, 1996, pp. 43–46), but without the requirement of specifying the Lipschitz constant. The algorithm also requires no knowledge of the objective function gradient, which makes it appealing in solving our problem where the effects of terrain and radio propagation make for a very complex function. The algorithm iteratively samples points within the domain, choosing these points based on the unexplored territory of the function and the previously calculated sample function values. The algorithm stops after a defined number of iterations.

He et al., (2004) use the DIRECT algorithm to solve an indoor wireless AP placement problem. They utilize very accurate but computationally expensive ray-tracing to calculate radio propagation, placing a premium on the number of required iterations. They conclude the DIRECT algorithm is an efficient means of finding good solutions relatively quickly.

The DIRECT algorithm is guaranteed to reach a global optimum if the objective function is continuous (Jones et al., 1993). To meet this requirement, we use bilinear interpolation to create a continuous solution space from our discrete map data.

### 1. Bilinear Interpolation

Bilinear interpolation is a common image processing technique that utilizes the four nearest neighboring points to calculate the intensity (in this case, elevation) of a given point (Gonzalez and Woods, 2007, p. 66). The technique provides adequately smooth elevation information, without the added computational load of bicubic interpolation or splines.

Recall we assume elevation $E$ is a function of location $uv$. We desire to find the elevation of a point $E_{xy}$, indicated by a double circle in Figure 9. The four nearest known points are $(x_1, y_1)$, $(x_1, y_2)$, $(x_2, y_1)$, and $(x_2, y_2)$, indicated by open circles. Their corresponding function values (elevations) are $E_{x_1y_1}, E_{x_1y_2}, E_{x_2y_1}, E_{x_2y_2}$.

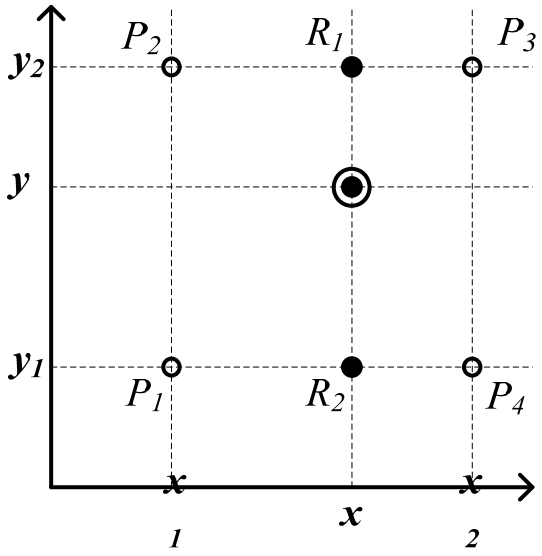


***Figure 9: Example of bilinear interpolation***

We first interpolate in the $x$ direction, to find the heights of the intermediate points $R_1$ and $R_2$, located at $(x, y_2)$ and $(x, y_1)$:

$$E_{xy_2} = \frac{x_2 - x}{x_2 - x_1} E_{x_1 y_2} + \frac{x - x_1}{x_2 - x_1} E_{x_2 y_2}$$

$$E_{xy_1} = \frac{x_2 - x}{x_2 - x_1} E_{x_1 y_1} + \frac{x - x_1}{x_2 - x_1} E_{x_2 y_1} .$$

We then interpolate in the *y* direction to determine the height of the desired interpolant *Q*:

$$E_{xy} = \frac{y_2 - y}{y_2 - y_1} E_{xy_1} + \frac{y - y_1}{y_2 - y_1} E_{xy_2} .$$

Replacement yields:

$$\begin{aligned} E_{xy} = & \frac{(x_2 - x)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} E_{x_1 y_1} \\ & + \frac{(x_2 - x)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} E_{x_1 y_2} \\ & + \frac{(x - x_1)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} E_{x_2 y_2} \\ & + \frac{(x - x_1)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} E_{x_2 y_1} . \end{aligned}$$

## 2. Theoretical Basis of DIRECT

Jones et al., (1993) build on Lipschitzian optimization to create the DIRECT algorithm. Essential to Lipschitzian optimization is the assumption that there exists a finite positive constant *k* that bounds the maximum rate of change of the function. That is,

$$\left| f(x) - f(x') \right| \le k \left| x - x' \right| \quad \forall x, x' \in [s, t]$$

where [*s*, *t*] are bounds within the domain of *f*, the objective function to be minimized. With this assumption, *k* can provide a lower bound on the minimum of *f* within [*s*, *t*]. Figure 10 demonstrates this graphically (many of the following figures are based on or embellish figures presented by Jones et al., 1993).

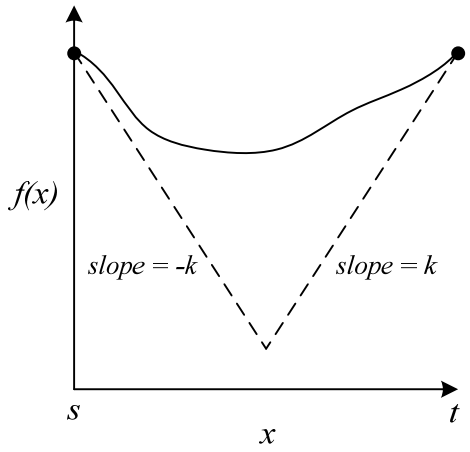


***Figure 10: Interval sampling of Lipschitzian optimization***

The minimum value of $f(x)$ within $[s, t]$ must be above the two dashed lines formed by the given maximum rate of change $k$ intersected at $s$ and $t$, respectively.

The Shubert (1972) algorithm raises the lower bound on the function by iteratively dividing the search space. Let $b$ denote the $x$ location where the two lines of slope $k$ intersect, and let $f(b)$ denote the value of the function at this point (see Figure 11).

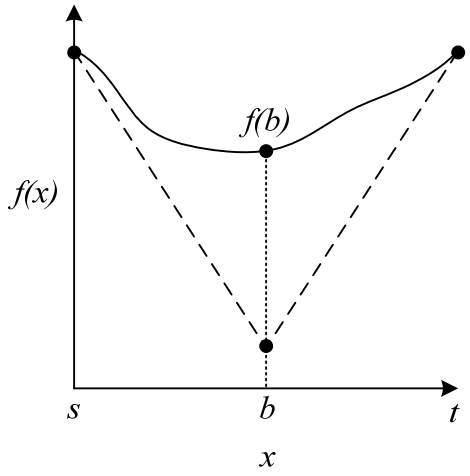


***Figure 11: First sample in Shubert algorithm***

The search space can be divided by creating two intervals [*s, b*] and [*b, t*], with new associated lower bounds and *f*(*b)* values based on *k*. The partition with the smallest *f*(*b*) is selected for the next iteration (ties are broken arbitrarily). This newly selected partition is again split into two, and the process repeats until the specified stopping conditions are met. The process is graphically demonstrated in Figure 12. In Figure 12a, the initial lower bound region is split into two at *b*. The left region is chosen to be split in Figure 12b, yielding two new lower bound regions. The process continues in Figure 12c.

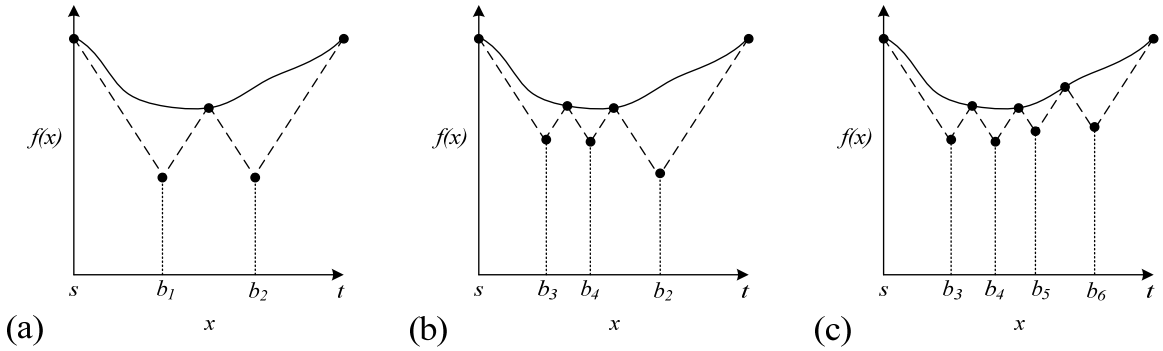


***Figure 12: Illustration of Shubert algorithm***

Jones et al., (1993) discuss two problems with the Shubert method. The first is that the rate of convergence is a function of the value of *k*, and this value is generally large. A larger *k* results in a larger area underneath a particular interval, causing a more global search and slower convergence. A smaller *k* searches locally, but may converge to a local minimum if the value is too small. Hence, the parameter must be carefully chosen to balance global and local search scope. The value can be reduced during the course of the algorithm, but it is not immediately apparent when it is appropriate to do so.

The second problem relates to the sampling method. The upper and lower endpoints of each dimension of a given interval must be sampled to calculate a lower bound. In the one-dimensional examples above only two samples are required, but this sampling per interval grows linearly with the number of dimensions. This quickly becomes computationally expensive, as the objective function is often complex and time-consuming to calculate. The DIRECT algorithm of Jones et al., (1993) addresses both of these problems.

### 3. Center-Point Sampling

To alleviate the complexities of sampling from each end of each dimension for a given interval, Jones et al., (1993) sample from the center of a given interval. Thus, the number of required sampling points per interval does not grow with the number of dimensions. Figure 13 demonstrates this in one dimension for a given *k*.

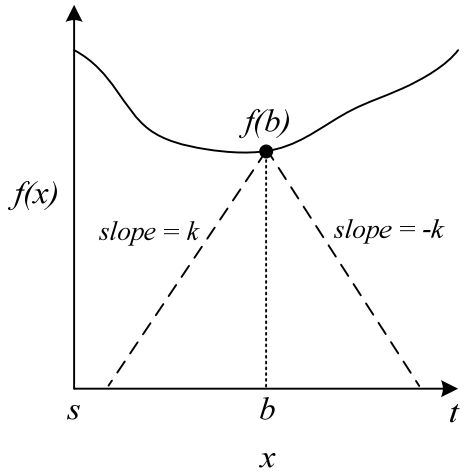


***Figure 13: Illustration of center-point sampling***

### 4. Rate of Change Constant

To overcome the problem of a fixed *k*, Jones et al., (1993) consider the objective value of a particular interval versus the interval's dimension size. Let $f(c_i)$ denote the function value of a particular interval $i$ at the center point $c$, and let $d_i$ denote the distance from the center point to the edge of an interval's dimension (i.e., $d_i = (t_i - s_i)/2$). Figure 14 plots $f(c_i)$ versus $d_i$ for various intervals (represented as dots).

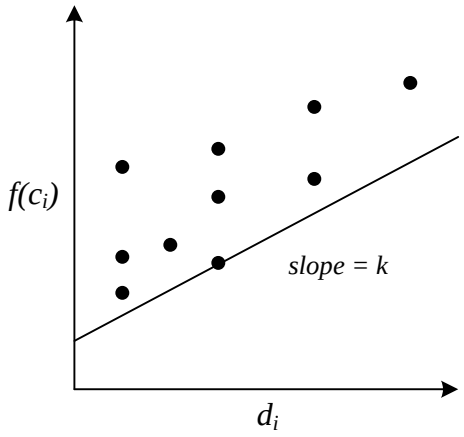


***Figure 14: Interval function value versus size***

Jones et al., (1993) observe that the vertical axis represents goodness with respect to local search: lower $f(c_i)$ values are closer to the actual global minimum. The horizontal axis represents goodness with respect to global search: larger $d_i$ values have more unexplored territory in their respective intervals. The lower bound for any interval can be found by intercepting a line of slope $k$ through the point of the desired interval with the vertical axis.

Jones et al., (1993) use this concept of goodness to introduce a rate-of-change parameter $\tilde{k}$ to vary the scope of search. The $\tilde{k}$ value can be changed during the course of the algorithm to find those intervals which are potentially optimal. Jones et al., (1993) define an interval $i$ to be potentially optimal if there exists some $\tilde{k} > 0$ such that

$$f(c_i) - \tilde{k}d_i \le f(c_j) - \tilde{k}d_j \quad \forall j = 1,...m \tag{16}$$

$$f(c_i) - \tilde{k}d_i \le f_{\min} - \varepsilon |f_{\min}| \tag{17}$$

where $f_{min}$ is the current best value, and $\varepsilon$ is a small positive constant. Equation (16) ensures that only those intervals on the lower-right of the convex hull of intervals can be potentially optimal. Equation (17) ensures that potentially optimal intervals must exceed

the current best objective value by a nontrivial amount (otherwise the search becomes too local). Jones et al., (1993) show that $\varepsilon$ values between $10^{-3}$ and $10^{-7}$ yield good results; we use $10^{-4}$.

Intervals that meet these two conditions are displayed as open dots in Figure 15. The value of $\tilde{k}$ defines the slope of the line. Notice that potentially optimal intervals must be on the lower-right of the convex hull of the cloud of intervals.

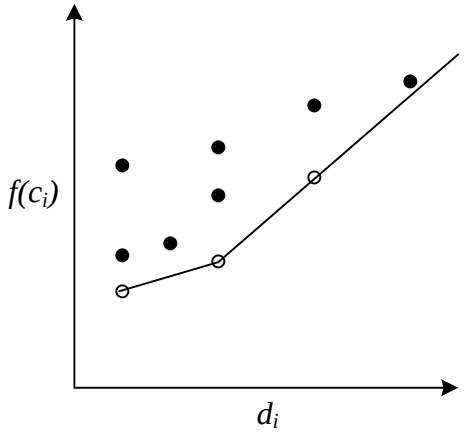


***Figure 15: Convex hull of potentially optimal intervals***

## 5. Finding Potentially Optimal Intervals

The next step is to efficiently find the set $\Pi$ of potentially optimal intervals. Jones et al., (1993) observe that since intervals are always divided into thirds, the only possible interval lengths are $(t - s)3^{-k}$ for $k = 0, 1, 2,\ldots$ Hence, many intervals have the same abscissa, as is evident in Figure 15. We use this key observation to create our algorithm.

Let $\Delta_i$ be the set of all intervals with largest dimension equal to $d_i$, that is, the set of all intervals with the same abscissas (see Figure 16a). Let $i$ =1, 2, …$D$ denote a particular $\Delta$ or $d$, where $D$ is the largest interval (the original interval). Let $\Phi(\Delta_i)$ be the element of $\Delta_i$ with the smallest objective value $f(c)$, that is, the element closest to the horizontal axis in Figure 16b. If there is more than one $\Phi(\Delta_i)$ for any interval length, we note this and evaluate any of them as representative of the subset.

We calculate the angle created between $f_{min}$ and each $\Phi(\Delta_i)$ for all $i$ with $d$ values larger than that of $f_{min}$ (see Figure 16c), and then create a set $S$ containing all $\Phi(\Delta_i)$, sorted by this angle from smallest to largest. We only consider the elements of set $S$ as extreme points in forming the convex hull specified by the potential optimality conditions.

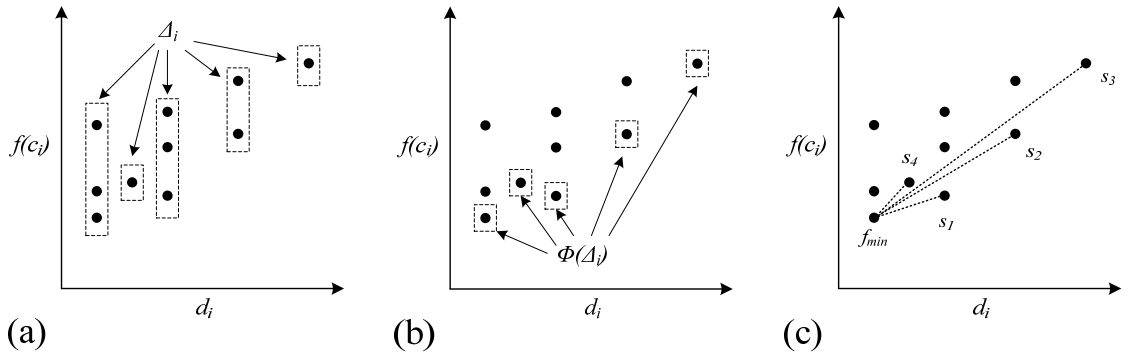


***Figure 16: Identifying Φ(Δi) for all intervals***

Conceptually, the algorithm begins at $f_{min}$ and adds it to the set $\Pi$. The first element $s_1$ of $S$ (that is, the $\Phi(\Delta_i)$ with the smallest angle from $f_{min}$) is selected, and a line of slope $\tilde{k}_1$ is drawn between the two points (see Figure 17a). The element $s_1$ is added to $\Pi$ if it meets the second potential optimality condition. If multiple $\Phi(\Delta_i)$ exist at a particular interval length, all are added to $\Pi$ if any of them satisfies the potential optimality conditions. All elements of $S$ with $d_i$ distances less than or equal to that of $s_1$ are removed (in Figure 17a, those elements to the left of $s_1$), and a line between $s_1$ and next element of $S$, $s_2$ is then drawn (see Figure 17b). The process repeats until the set $S$ has been exhausted (see Figure 17c).

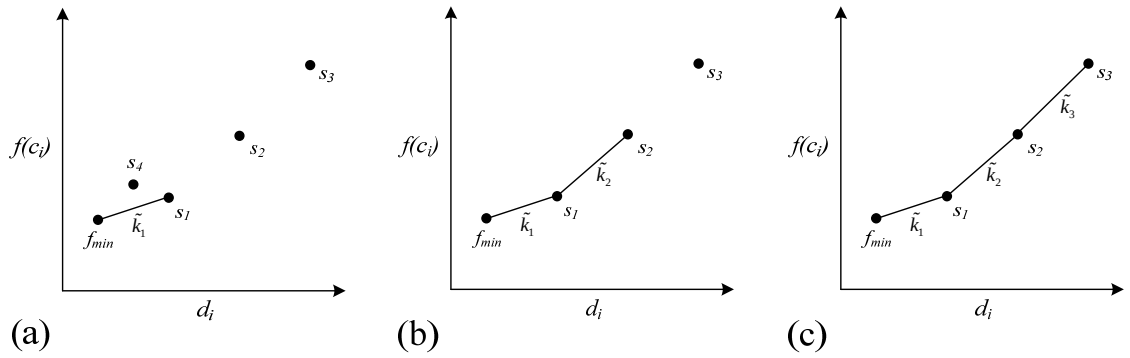


***Figure 17: Calculating $\tilde{k}_i$ and potentially optimal intervals***

The following algorithm computes *Π*, the set of potentially optimal intervals.

**Algorithm Potentially Optimal**

Input: $S$, $\Phi(\Delta_i)$, $f(\Phi(\Delta_i))$, and $d_i$ for all $i = 1, 2, \ldots D$, $\varepsilon$ value, $f_{min}$

Output: Set *Π* of potentially optimal intervals

**begin**
Set *place* = $\Phi(\Delta_1)$
Add *place* to *Π*
**for** ( $\forall s \in S$ )
Set $\tilde{k}_i = [f(s) - f(place)] / (d_s - d_{place})$
**if** $(f(s) - \tilde{k}_i d_s \leq f_{min} - \varepsilon|f_{min}|$
Add *s* to *Π*
**end if;**
Remove *s* from *S*
Remove any elements of *S* with $d_i \leq d_s$
place = *s*
**next** s**;**
**end;**

### 6. The DIRECT Algorithm

Without loss of generality, Jones et al., (1993) extend the one-dimensional process to multiple dimensions. Rather than just splitting a line into smaller segments (as in the one-dimensional case), the DIRECT algorithm iteratively divides the multidimensional solution space into smaller hyper-rectangles (hence the name DIviding RECTangles), based on their potential optimality.

To initialize the algorithm, each dimension of the solution space is normalized to have a lower bound of zero and an upper bound of one, creating a unit hyper-cube. The algorithm evaluates the function at the center of the hyper-cube. Next, the algorithm determines the set $\Pi$ of potentially optimal hyper-rectangles/cubes. During the first iteration, there is only one hyper-cube. The algorithm then samples each dimension of each element of $\Pi$ at one-third the side-lengths of the largest dimension (in the case of a hyper-cube, all dimensions are sampled). The algorithm determines these locations by calculating $c \pm \delta e_i$, where $c$ is the center point, $\delta$ is one-third the side-length of the largest dimension, and $e_i$ is the $i^{th}$ unit vector. We demonstrate this for two dimensions in Figure 18.

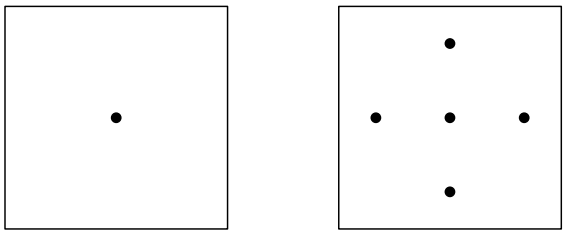


***Figure 18: Sampling the initial hyper-cube***

In Figure 18a, the algorithm samples the center of the hyper-cube. In Figure 18b, the algorithm samples along each of the hyper-cube's largest dimensions (i.e., vertical and horizontal dimensions). Since the interval is a hyper-cube with equal length dimensions, the algorithm samples all dimensions.

Next, the algorithm computes $w_i$ (the smallest sampled value along each dimension) for each element of $\Pi$:

$$w_i = \min\left\{ f\left(c + \delta e_i\right), f\left(c - \delta e_i\right)\right\} \quad \forall i \in Q \tag{18}$$

where $Q$ is the set of the largest dimensions of a particular hyper-rectangle/cube, and $e_i$ is the $i$th unit vector. The dimension $i$ with the smallest $w_i$ value contains the best function value. Jones et al., (1993) place this point in the largest sub-interval by first dividing the

original hyper-rectangle/cube along the dimension with the smallest $w_i$. The process continues for the dimension with the next smallest $w_i$ value (see Figure 19).

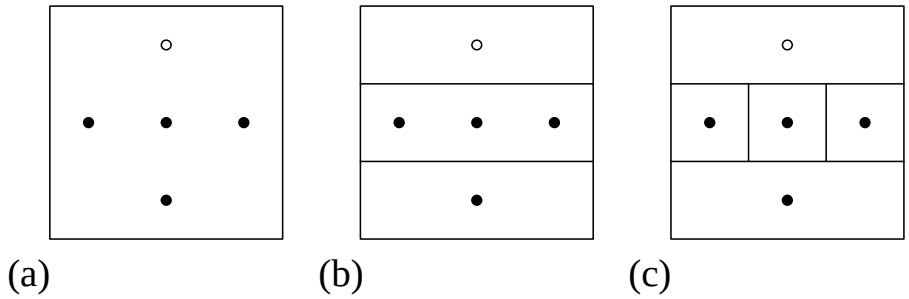


***Figure 19: Division of the initial hyper-cube***

In Figure 19a, the top sampled point (open dot) has the best objective value, so the vertical dimension contains the smallest $w_i$ value. In Figure 19b, this dimension is divided first, and in Figure 19c, the horizontal dimension is divided. The sampled points then become the centers of the new hyper-rectangles/cubes.

Following the division, the value of $f_{min}$ is updated, and the set $\Pi$ of potentially optimal hyper-rectangles/cubes is recalculated. The process continues until a predefined number of iterations or function evaluations has been reached.

Following Jones et al., (1993), the below pseudo-code outlines the DIviding RECTangles algorithm.

**Algorithm DIRECT**

Input: objective function *f*, number and size of dimensions of solution space, $\varepsilon$ value, maximum number of DIRECT iterations and function evaluations

Output: Value and location of best solution found

**begin**
    Normalize solution space
    Calculate center *c* of hyper-cube
    Evaluate *f* at $c \pm \delta e_i$ for each dimension
    Set *min_location* = center of hyper-cube
    Set $f_{min} = f(min\_location)$
    **while** (*iteration* < *max_iterations* and *evaluations* < *max_evaluations*) **do**
        **if** (*iteration* = 1)
            Add $\Phi(\Delta_1)$ to set $\Pi$
        **else**
            Calculate set *S* of all $\Phi(\Delta_i)$
            Identify set $\Pi$ using *Algorithm Potentially Optimal*
        **end if;**
        **for** (all $j \in \Pi$ )
            Identify longest side(s) of hyper-rectangle/cube *j*
            Evaluate *f* at $c \pm \delta e_i$ for each longest dimension
            Increase *evaluations*
            Calculate $w_i$
            Divide *j* into smaller hyper-rectangles/cubes based on $w_i$
            Update $f_{min}$ and *min_location*
        **next j**
        Increase *iteration*
    **end;**
    Convert *min_location* to original coordinate system
**end;**

### *a. Convergence*

Jones et al., (1993) show that DIRECT is guaranteed to converge to the global optimum if the objective function is continuous. The basis of the proof relies on the fact that as the number of iterations goes to infinity, the size of the hyper-rectangles goes to zero, so DIRECT will eventually sample a point within an arbitrary distance of any desired point.

One drawback to the DIRECT algorithm is that the distance to the global optimum at any iteration cannot be accurately calculated. If the Lipschitz constant *k* is known, a lower bound can be calculated by evaluating the function at the boundaries of each dimension, drawing a line of slope *k* through each evaluated point, and intersecting the lines (similar to Figure 11). In our application, finding this constant would be as difficult as solving the problem to optimality.

As noted in Chapter III.D.2, we can calculate a theoretical (but likely infeasible) lower bound to the SRRA+C objective value. We determine the best possible network flow value by simply calculating network flow with virtually no propagation loss between APs. We then add the best possible coverage shortfall value (which is zero), and the result is a lower bound to the function. The value is likely infeasible, but this method can at least provide some metric to gauge the progress of the DIRECT algorithm. We examine this in further detail in Chapter IV.B.4.

#### *b. Performance and Limitations*

Following Jones et al., (1993) and Finkel (2003), we assess our implementation of DIRECT using several test functions, including the Six-Hump Cambelback (two dimensions), Branin Function (two dimensions) Shekel S5 (four dimensions), and Hartman H6 (six dimensions). We find our implementation to yield similar (but not identical) performance results. Specifically, we find our implementation requires more function evaluations to obtain the same level of relative error. This disparity may be due to a difference in counting techniques, or computational inefficiencies on our part.

The limit of precision in our implementation of the DIRECT algorithm is determined by our method of calculating the center of each hyper-rectangle. We store the positions of each boundary for each hyper-rectangle, and compute the center point by finding the middle of these boundaries. When DIRECT can no longer discern differences in the side lengths, it is unable to accurately find the center position, and hence cannot correctly evaluate the function value.

To determine this computational limit, we note that the side length of any hyper-rectangle is $3^{-d}$, where $d$ is the number of times that side has been divided. The limit occurs when, given the floating point precision of the storage method, $3^{-d}$ is indistinguishable from $3^{-d+1}$. The double-precision floating point variables of Visual C++ are precise to approximately 15 digits (Microsoft, 2009). Simple calculation shows this limit occurs at approximately $d$ = 32 divisions. While some languages (such as FORTRAN) are able to handle quadruple-precision floating point variables and hence many more divisions, we have not found this limitation of C++ to impair our ability in any way to find good solutions with DIRECT. Given a coverage region of one hundred square kilometers, 32 divisions would yield an interval side length of approximately 5.39 x $10^{-9}$ millimeters!

## F. SRRA+C WITH DIRECT

We use the DIRECT algorithm iteratively to choose AP locations for solving the SRRA+C problem. The solution space is defined by the operating area in the form of terrain data. The DIRECT algorithm requires this area to be in the shape of a rectangle or square, to facilitate transformation into a unit hyper-cube.

The dimensionality of the problem follows from the number of APs that can move, ($n$-1). Each moving (i.e., non-HQ) AP has an associated ($u$, $v$) location bounded by the operating region, so both of these dimensions must be considered for each AP. Hence, the number of dimensions is 2($n$-1). A point in this solution space represents ($u$, $v$) coordinates for every AP.

### 1. Hierarchical Decomposition

Figure 20 summarizes the hierarchal method of solving the SRRA+C problem. For each iteration of DIRECT, the algorithm chooses a set of AP locations $\sigma$. We compute coverage shortfall and the SRRA solution for these locations, and we combine the values into the overall objective function value $\Omega(\sigma)$. The process continues until we reach the desired stopping criterion. Possible criteria include elapsed time, size of

interval side length, and number of iterations since last objective value improvement or change in AP locations. We use a desired number of DIRECT iterations.

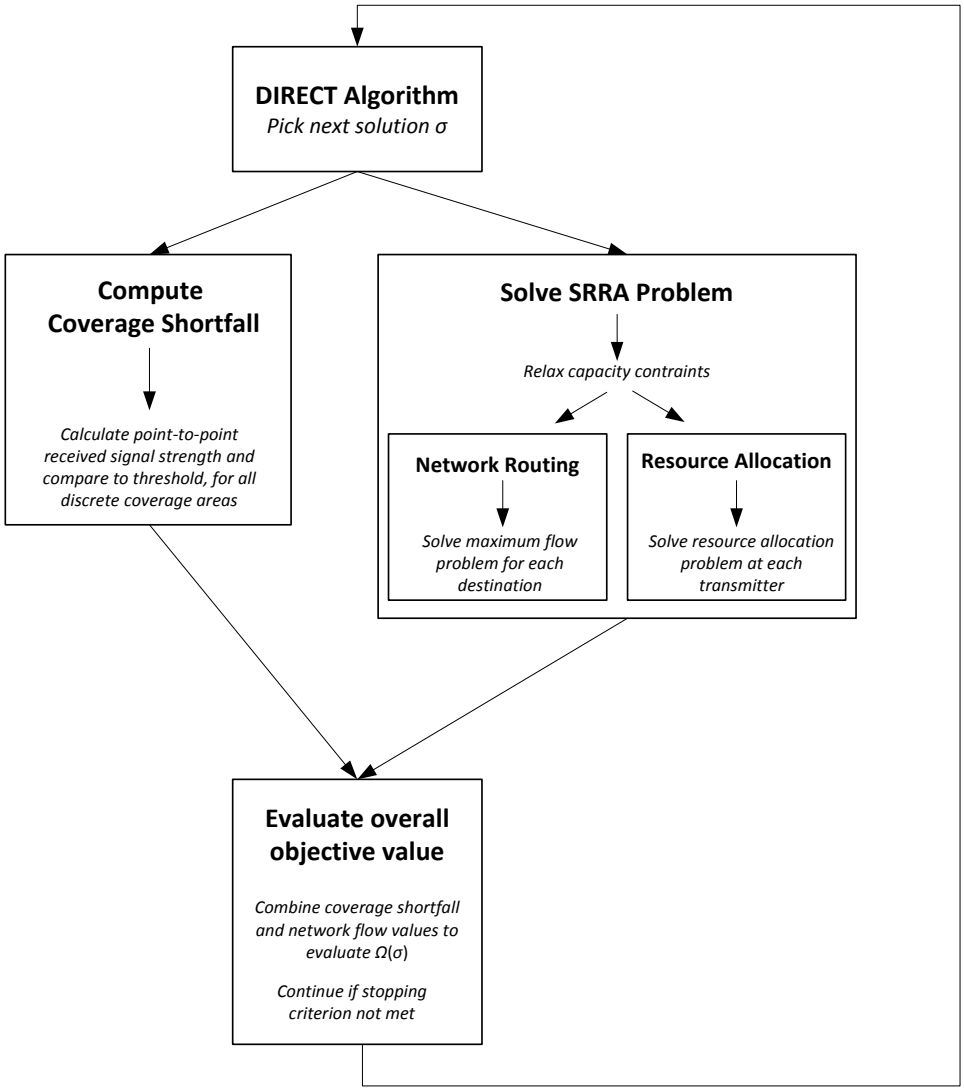


***Figure 20: Hierarchical decomposition of SRRA+C solved with DIRECT***

### 2. SRRA+C with DIRECT Algorithm

The following pseudo-code outlines the SRRA+C process utilizing the DIRECT algorithm.

**Algorithm SRRA+C with DIRECT**

Input: Number and characteristics of APs, desired coverage region, HQ node location, minimum flow requirement, map data, DIRECT stopping criteria, and ε value (for DIRECT)

Output: WMN value; location, traffic routing and power allocations for all APs; client received signal strength at each discrete map location (*u*,*v*)

**begin**
    Store map data
    **while** (DIRECT stopping criterion not true) **do**
        Calculate AP locations using *Algorithm DIRECT*
        Calculate coverage shortfall value
        Calculate network flow values using *Algorithm SRRA*
        Calculate overall objective value
    **end;**
**end;**

## G. SRRA+C DECISION SUPPORT TOOL

To solve the SRRA+C problem, we created a graphical user interface (GUI) in Microsoft Visual C++ (Microsoft Corporation, 2009). The stand-alone program reads text files in a generic XYZ format as input for terrain elevations. This map information can be of any scale and any grid-based format (e.g., Universal Transverse Mercator (UTM) coordinate system). The tool allows the user to change all input data described in our formulation. See Appendix A for a complete list of inputs and outputs. A specific problem and its associated solution can be saved as text files for later use.

The program allows three different modes of radio propagation (inverse-square, Hata COST-231, and TIREM). The TIREM model is accessed via a dynamic link library (DLL) provided by the U.S. Joint Spectrum Center. The program can solve the SRRA+C problem using either enumeration or the DIRECT algorithm.

The program has two operating modes. The first is a Coverage Analysis mode, used to solve the SRRA+C problem as described. The user is able to input all required variables, including drawing the coverage region directly on a graphical display of the terrain map. Figure 21 is a screenshot of our decision support tool in Coverage Analysis mode. The central panel displays the terrain of the operating region (green/darker area is lower elevation, orange/lighter area is higher elevation). The box on the terrain indicates the user-defined client coverage region. The right panels contain controls for all input data.

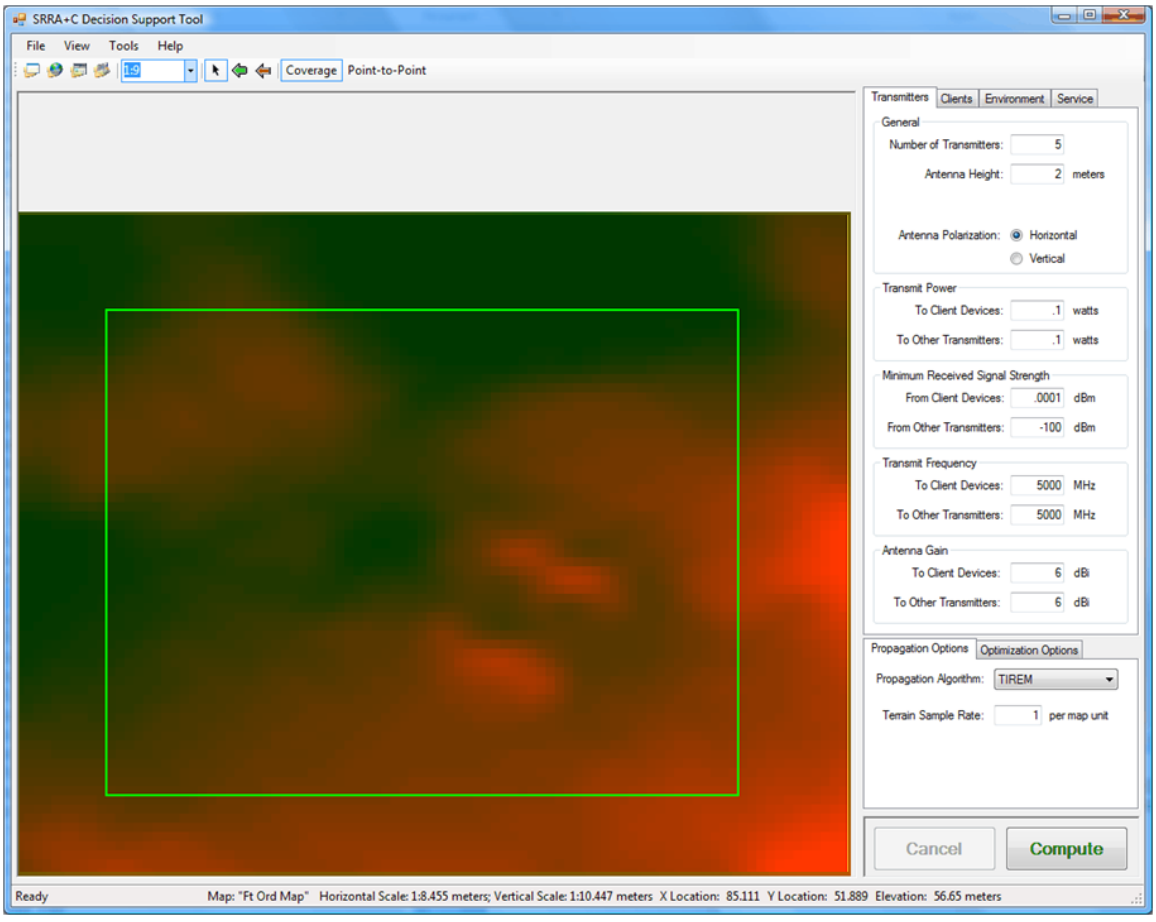


***Figure 21: Screenshot of decision support tool in Coverage Analysis Mode, displaying terrain and coverage region*** **[color online]**

After solving the problem, the programs displays the locations of APs and their associated links, and the areas in the coverage region receiving adequate client coverage graphically on the map. Figure 22 is a screenshot of a completed SRRA+C optimization. AP locations are indicated by transmitter icons, and calculated traffic flow links are indicated by lines connecting APs. Within the defined coverage region, green/lighter area indicates adequate coverage and red/darker area indicates coverage shortfall.

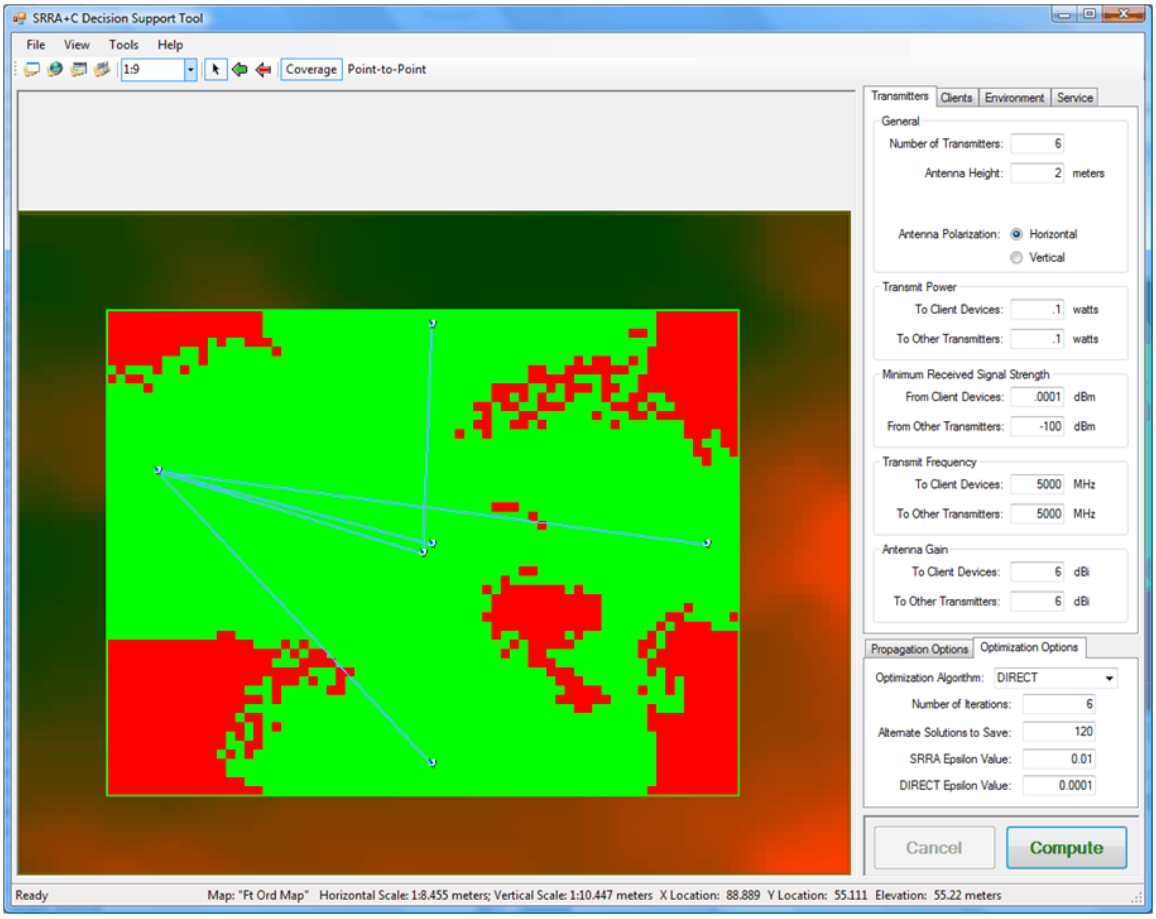


***Figure 22:*** ***Screenshot of completed SRRA+C optimization, indicating network topology and coverage shortfall*** **[color online]**

The other operating mode is a Point-to-Point Analysis mode. This mode is used to compute point-to-point losses, received signal strengths, and theoretical link capacities between any two points. A side-view of the terrain profile, as well as LOS paths and the first Fresnel zone, are displayed graphically by the tool (see Figure 23). This analysis mode is useful for examining proposed site locations, as well as tuning the inputs of the Coverage Analysis mode to provide desired output.

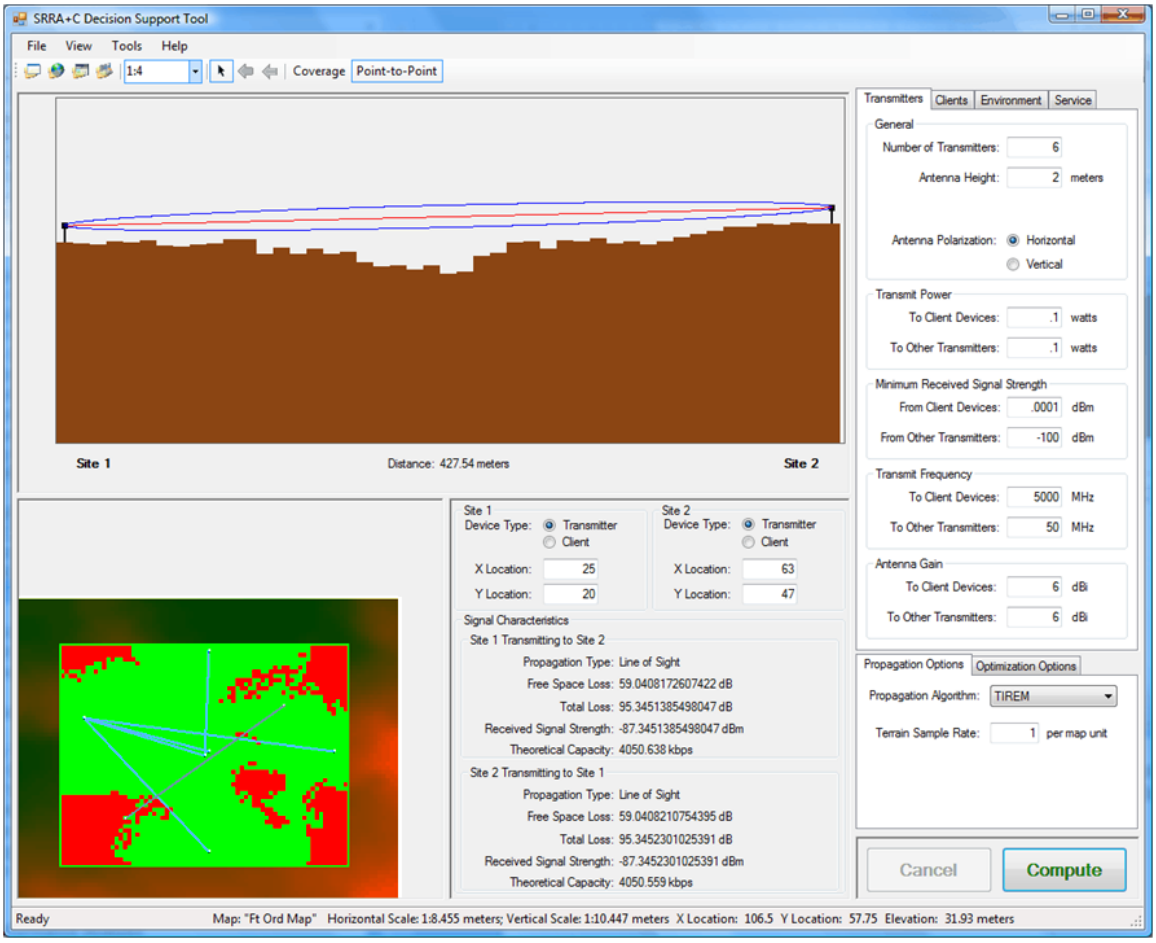


***Figure 23: Screenshot of decision support tool in Point-to-Point Analysis mode, displaying path propagation and elevation profile*** **[color online]**

# IV. ANALYSIS AND RESULTS

Our analysis of the SRRA+C formulation and solution methods begins with field testing to validate to each portion of the objective function, coverage and network flow. We then conduct an analysis of the formulation by examining examples of the solution surface and considering the effects of two design variables. Next, we use SRRA+C to solve small problems to gain an understanding of its operation. We also compare the performance of the DIRECT solution method to total enumeration. We then present the results of our network topology field tests using SRRA+C with DIRECT. We conclude with general observations of our formulation and the DIRECT solution method.

## A. SRRA+C MODEL VALIDATION

The validity of the SRRA+C formulation, specifically its ability to quantify network performance, rests on the accuracy of radio wave propagation and network throughput predictions. Each part of the overall objective function depends on the algorithms associated with these predictions. The client coverage portion of the formulation relies directly on received signal strength, a function of radio wave propagation. Network throughput depends on received signal strength and the Shannon capacity formula. We conduct simple field tests to provide some level of validation for each algorithm.

### 1. Validation of Client Coverage Model: Received Signal Strength Test

We use received signal strength to calculate both client coverage and the traffic capacity between APs. While the Hata COST–231 model can provide a decent approximation of radio propagation, the criticality of accuracy inclines us to use TIREM for all tests. Though TIREM is computationally more expensive than COST–231, in practice we find the runtime differences to be negligible.

As discussed in Chapter III, TIREM predicts radio propagation between two points, given a profile of elevations between the points. The model has proven its value

in many government and commercial modeling applications, but it does have its limitations. It does not account for foliage, rain, man-made obstacles, and any obstacles not directly in the path between the two given points. Nor does it account for any long-term variability; the predictions are strictly deterministic. For line-of-sight calculations, the difference between TIREM predictions and actual measurements has been show to have a mean of −2.8 dB and a standard deviation of 8.9 dB (Eppink and Kuebler, 1994). Considering all the variables contributing to propagation physics, this level of variation is very reasonable.

To provide further validation of TIREM, we measure this difference using the equipment we utilize during the full network field test, over the actual terrain. We conduct a point-to-point test between a fixed Cisco AP1000-series Aironet WMN AP (Cisco, 2009) and an Intel wireless transceiver internal to a mobile laptop computer. We measure received signal strength (RSS) in dBm as a function of distance in meters. The respective AP and client device gains are 6 dBi and 3.5 dBi, and the respective antenna heights are two meters and one meter. Transmitting power is approximately 20 dBm. We conduct the test aboard Fort Ord, California over generally flat terrain consisting of pavement and packed gravel, with no trees or other obstructions to the LOS path.

The first test is conducted at the 5.8 GHz operating frequency. We collect $n$ = 200 observations from 0 to 466 meters. Using a fade margin of 10 dB, we compute the differences between TIREM predictions and actual observations, and find the mean of these differences to be −0.64 dB and the standard deviation to be 4.97. Figure 24 illustrates the results of this field test.

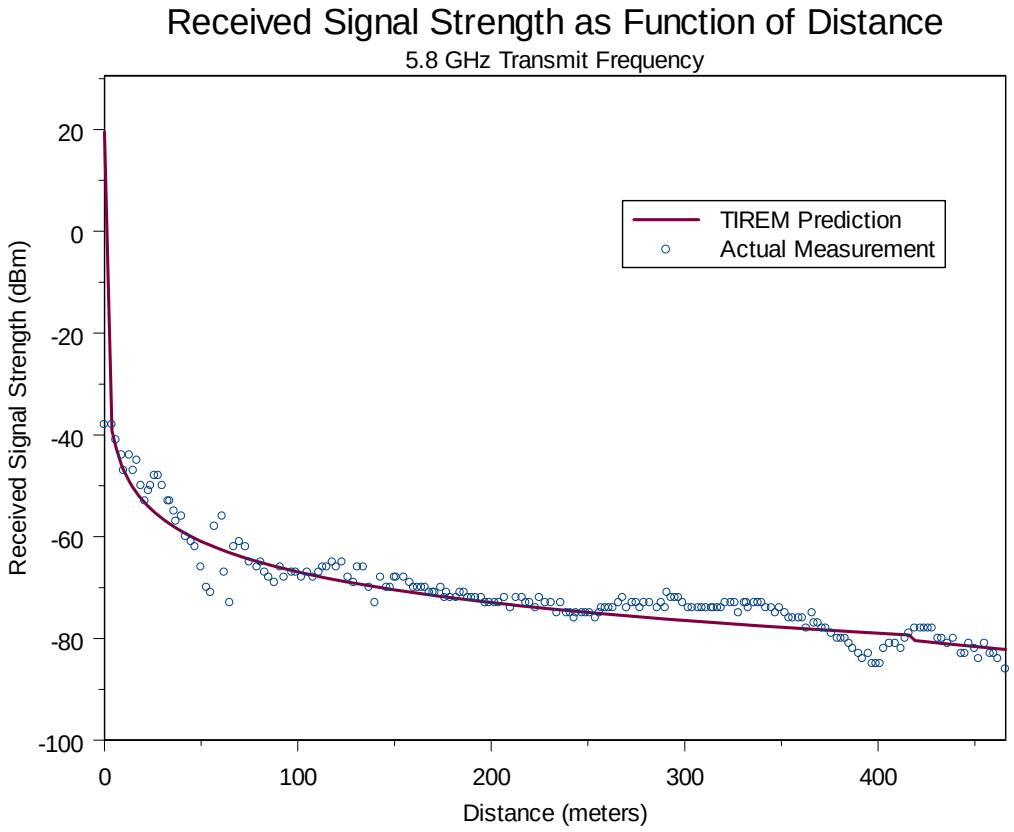


***Figure 24: Results of point-to-point field test of received signal strength at 5.8 GHz***

We conduct a similar test at the 2.4 GHz operating frequency. We collect $n$ = 165 observations from 0 to 464 meters. We find the mean of the differences to be −2.65 dB and the standard deviation to be 7.56. Figure 25 illustrates the results of this field test.

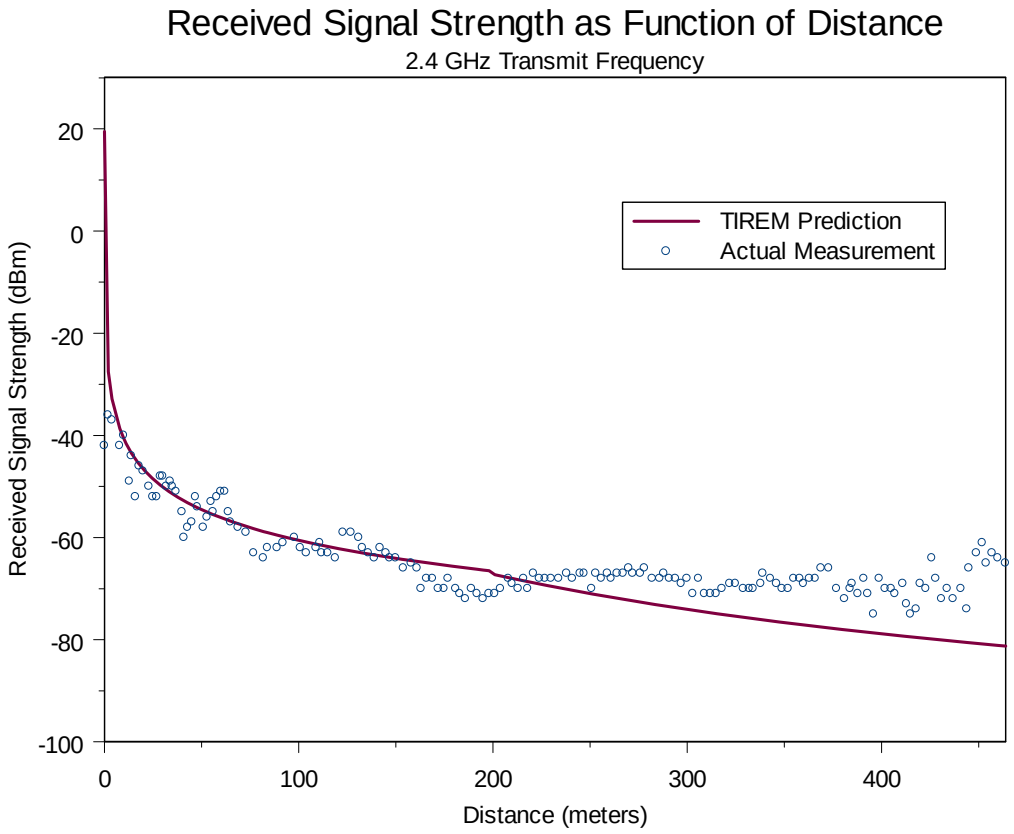


***Figure 25: Results of point-to-point field test of received signal strength at 2.4 GHz***

Note in this test the actual measurements from 300 meters onward depart from the predicted TIREM values. We believe this is due to the presence of nearby buildings at that end of our testing range: the radio waves may have reflected off the buildings and provided a stronger signal than would have otherwise been received. If this is the case, it demonstrates one of the weaknesses of the TIREM model.

These two tests demonstrate that TIREM is capable of making very reasonable received signal strength predictions using our testing equipment in a real-world

environment. Since our formulation of client coverage depends solely on received signal strength, we conclude this formulation is an acceptable method of valuing coverage.

### 2. Validation of Network Flow Model: Link Capacity Test

Our SRRA network flow model depends both on prediction of received signal strength and output of the Shannon formula to determine traffic capacity between any two nodes. Recall the Shannon formula predicts the upper bound on the amount of information that can be transmitted between two points. This theoretical limit implicitly assumes perfect signal modulation, and so with an accurate background noise value, throughput with real-world equipment is less than calculated by Shannon.

We conduct a point-to-point field test to compare actual and Shannon-calculated throughput. Utilizing the same equipment and range as the previous tests, we compute actual traffic flow between the mobile laptop computer and another laptop attached to the fixed WMN AP by transferring a very large file. We then use TIREM and the Shannon capacity formula with a background noise level of -88 dBm to calculate an upper bound on link throughput. We conduct the first test at 5.8 GHz, and the results are displayed in Figure 26:

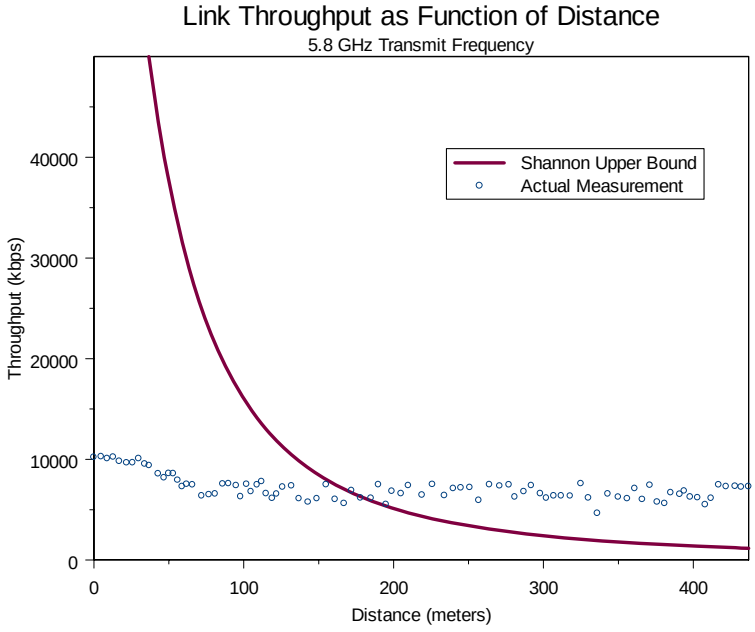


***Figure 26: Results of point-to-point field test of throughput at 5.8 GHz***

We then conduct the same test at 2.4 GHz:

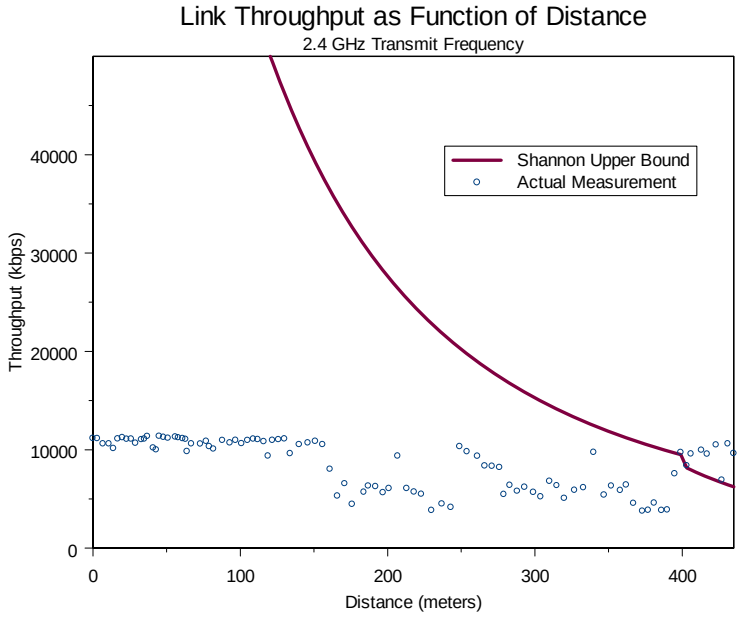


***Figure 27: Results of point-to-point field test of throughput at 2.4 GHz***

Both tests demonstrate the difficulties of accurately predicting throughput. While the previous field test demonstrates that TIREM can approximate actual received signal strength very well, other factors are clearly contributing to actual throughput. These may include the effects of proprietary network algorithms, hardware incongruities, or aspects of the transferred file itself. The single most important factor, however, is probably variation in background noise level. The test range is located near the campus of California State University Monterey Bay, and other wireless networks (as well as a host of other radio-emitting devices, such as cordless phones) that cause a varying level of background noise (see Figure 45, the results of our background noise analysis in this area). In Figures 26 and 27, a larger background noise level would lower the Shannon upper bound.

Note the sharp drop in the Shannon upper bound near 400 meters in Figure 27. These drops are present but less visible at different distances in Figures 24 and 25. TIREM is a conglomeration of many different propagation models, and each model is

used for certain input values. We believe these drops indicate "handoff" of received signal strength from one predictive model to the next as a result of varying distance.

## B. SRRA+C ANALYSIS

We next conduct analysis on the SRRA+C solution space. This provides some indication of the robustness of the formulation to varying environments and topology scenarios, as well as how likely it may be for DIRECT to find good solutions quickly. All instances of the use of our decision support tool in this chapter are conducted using a Windows XP laptop with an Intel Core 2 Duo processor operating at 2.16 GHz and 2 GB of RAM. Terrain information is provided by the United States Geological Survey (USGS) via MapMart (2009).

### 1. Rank-Ordered Solutions

We first examine the sensitivity of the objective value as a function of the rank of the solution. That is, we wish to determine how good the optimal and near-optimal solutions are in relation to the rest of the possible solutions. If the optimal solutions are significantly better than any other, then it may be difficult for DIRECT to find those optimal solutions. Recall from Chapter III that DIRECT samples the solution space using a set pattern. If the optimum appears very suddenly on the solution surface, DIRECT may not sample near enough to the optimal valley to discern its presence.

We conduct the analysis by using our decision support tool to enumerate all discrete solutions of a three AP network on Fort Ord terrain consisting of 65 x 33 grid locations. The analysis enumerates $\binom{(65)(33)}{3-1}$ = 2,299,400 unique solutions in eight hours, 38 minutes, and 55 seconds. We rank the solutions by objective value to produce the following graph:

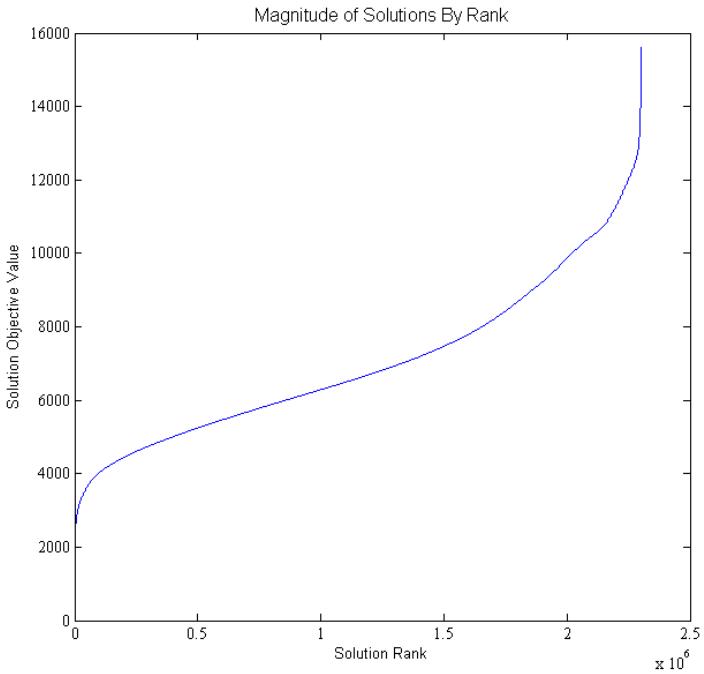


***Figure 28: Sensitivity analysis of SRRA+C objective value***

The graph shows a very small number of solutions that have very undesirable values (on the far right), a roughly linear progression to better solutions, and a very small number of solutions that are significantly more desirable (far left). This suggests that the surface of the SRRA+C solution space may have sharp troughs of optimality, along with very high peaks of bad solution values.

### 2. Solution Surface

We next create plots of the SRRA+C solution surface. Recall the number of dimensions of the SRRA+C problem is $2n-1$, where $n$ is the number of APs. To plot in three dimensions, we are limited to two independent dimensions and so may only plot the solution surface of a two AP network: one AP is the fixed HQ node, and the other moves in the X and Y directions (easterly and northerly).

We enumerate a sample two AP network over the Fort Ord terrain with the HQ node fixed at a central position. With an elevation map consisting of 65 x 33 discrete

coverage locations, there are $\binom{(65)(33)}{2-1}$ = 2,145 unique solutions. We plot the overall objective value as a function of the position of the moving AP; the results are presented in Figures 29 and 30. In Figure 29, the Z axis represents the objective value (lower is more desirable). Figure 30 is a contour plot of the solution surface, where darker areas represent more desirable solutions. The forward corner of Figure 29 is equivalent to the lower-left corner of Figure 30.

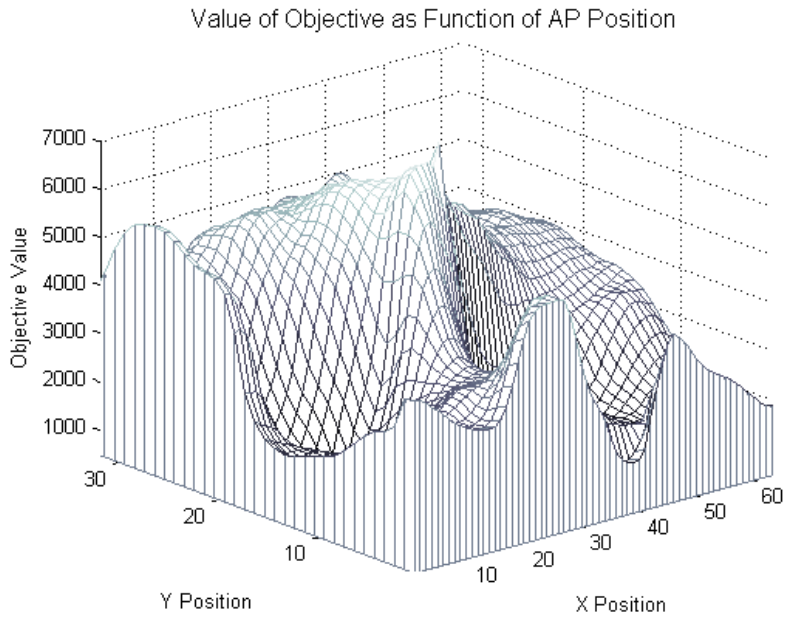


***Figure 29: Example of SRRA+C solution surface***

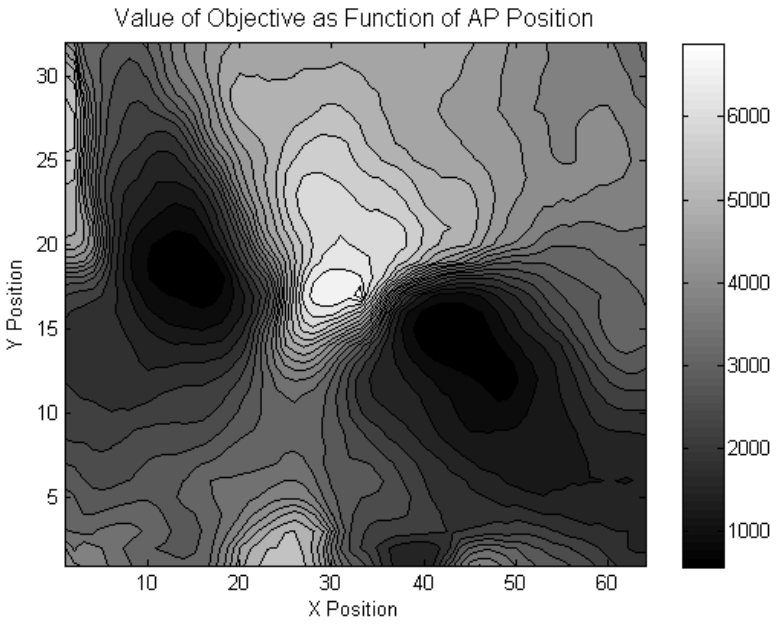


***Figure 30: Contour plot of example SRRA+C solution surface***

While it is impossible to create such figures for higher-dimensional problems, we can glean useful information from these plots. First, we note the solution space does appear to be continuous, which is a requirement for the convergence of the DIRECT algorithm. We also see further proof of the observations we drew from Figure 28: there appear to be many mediocre solutions (grayish areas), with a few sharp peaks of very undesirable solutions (white areas) and troughs of local optima (dark areas). The presence of multiple local optima suggest DIRECT may fall into a bad local optimum, but as described in Chapter III, DIRECT is guaranteed to eventually break out and find the global optimum as the number of iterations goes to infinity.

Additionally, by comparing the solution surface with the terrain elevation (see Figure 41), we see that points of higher elevation seem to be more desirable. This is in agreement with the general propagation principles described in Chapter II.

### 3. Sensitivity to Flow Value Weight *w*

Next, we examine the sensitivity of the solution to changes in *w*, the weight assigned to the network flow value in the overall objective function. Recall that a larger

value of *w* places greater emphasis on network flow. This should in general cause the network topology to be more compact, as this reduces the propagation losses between APs.

We test this by examining a network of five APs aboard Fort Ord using DIRECT and our decision support tool. We vary *w* while fixing all other variables, and run DIRECT for 10 iterations. We measure the sum of the distances between each AP and the HQ node, the client coverage shortfall values, and the total amount of flow delivered to the HQ AP in kilobits per second (kbps). The results are tabulated below.

| ***w*** | **Sum of Distances (meters)** | **Coverage Shortfall Value** | **Total Flow Delivered (kbps)** |
|---|---|---|---|
| 100 | 519.99 | 1651.98 | 17375.974 |
| 200 | 466.51 | 2044.905 | 23510.476 |
| 300 | 428.28 | 1946.461 | 23832.668 |
| 400 | 398.32 | 3005.367 | 142745.694 |
| 500 | 226.09 | 4991.718 | 149682.586 |
| 600 | 216.21 | 5327.257 | 143301.138 |
| 700 | 216.21 | 5327.257 | 143301.138 |
| 800 | 216.21 | 5327.257 | 143301.138 |
| 900 | 216.21 | 5327.257 | 143301.138 |
| 1000 | 222.59 | 6712.523 | 160520.929 |

***Table 2: Effects of varying flow weight w on SRRA+C output***

The results show that in general, a larger weight *w* increases coverage shortfall, and decreases the sum of the distances between the HQ node and all outlying nodes. Note that these generalizations do not always hold: coverage shortfall *decreases* when *w* is increased from 200 to 300, and the sum of distances *increases* when *w* is increased from 900 to 1000.

These nonmonotonicities result from our implementation of the subgradient method within the SRRA+C algorithm. Recall, that following Xiao et al., (2004), we use the subgradient method to solve the SRRA problem and determine optimal network flow for given AP locations. We run the subgradient method to a given number of iterations—not to optimality—and use the best dual objective value found as an estimate for network flow. The distance from convergence depends on the condition number of the problem (a

function of input values), since an ill-conditioned SRRA problem requires more iterations to converge. Hence, in some cases the actual flow estimate may be nonmonotonic in the input variables. The subgradient method could be run to optimality to overcome these small variations, but the computational cost would be prohibitive.

### 4. Theoretical Lower Bound

As noted in Chapters III.D.2 and III.E.6, we can calculate a theoretical lower bound on the overall objective value by solving for a "perfect" network: a solution with a coverage shortfall value of zero and the best possible network flow. While this solution may be infeasible, we consider an example to illustrate the value of this technique.

We examine a network of five APs over a 145 acre section of Fort Ord. We solve using DIRECT, and stop the algorithm at each iteration 1, 2,…30 to record the current overall objective value. We plot these values and compare to our theoretical lower bound for this network. The results are displayed in Figure 31.

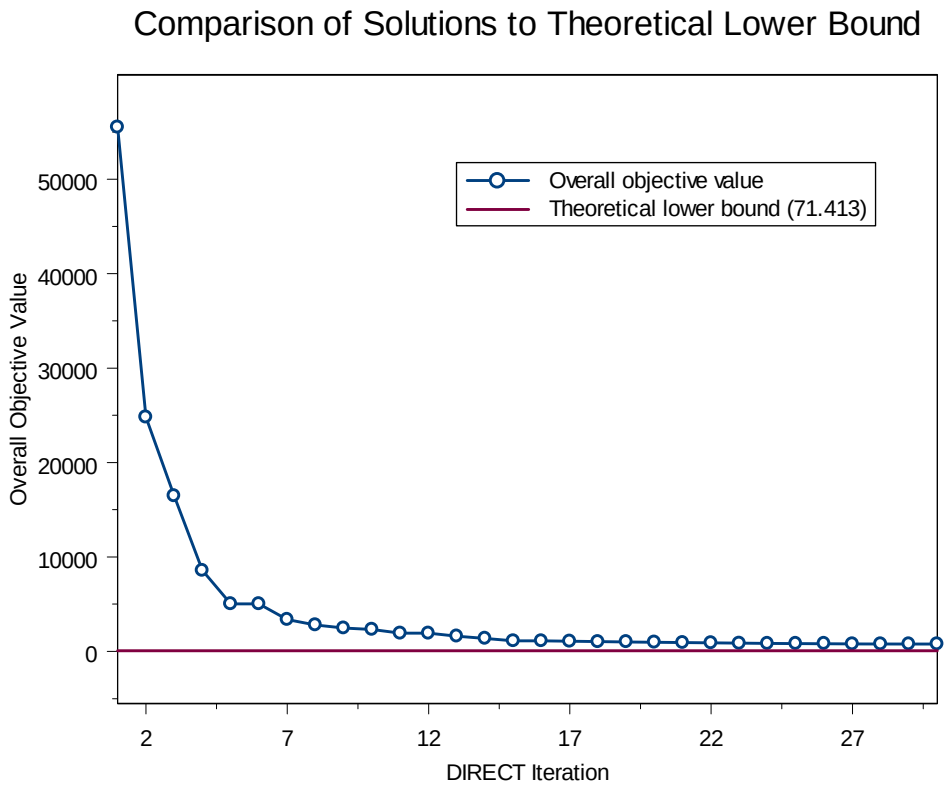


***Figure 31: Example comparing DIRECT solutions to a theoretical lower bound***

While the theoretical lower bound is infeasible in this problem, we see that the best solution value computed by DIRECT after 30 iterations (796.02) is within an order of magnitude of the value of the theoretical lower bound (71.413). We emphasize this is not a certificate on the optimality gap, but it does tell us we are relatively close to the global optimum (whatever it may actually be).

We compare the performances of DIRECT and enumeration in Chapter IV.C, but this example is particularly poignant in demonstrating the speed of DIRECT. To solve this problem to 22 iterations of DIRECT takes slightly more than 17 seconds; to solve this problem using our enumeration technique would take more than 47,290 *years!*

### 5. Pareto Frontier of Solutions by Number of APs

We now create a Pareto frontier balancing both terms of the overall objective function $\Omega$ (Equation 15) as a function of the number of APs. The information presented in such a Pareto frontier serves two purposes. First, like the theoretical lower bound discussed in the previous section, it provides information on the relative goodness of a particular solution $\sigma$ by comparing it to the bounds of network flow and coverage shortfall. It also can provide a quantification on the value of additional APs and the appropriate number of APs for a particular scenario.

We consider a 45 acre section of Fort Ord (the same we use in our field test in Chapter IV.D). We use DIRECT to solve for networks of 2,3,…8 APs with a $w$ value of one, running the algorithm until the computer runs out of available memory or we reach the computational limit of 32 hyper-rectangle divisions (whichever comes first). Networks with more APs have more dimensions and a corresponding greater number of sub-hyper-rectangles. Hence, in general for larger networks, the DIRECT algorithm must be run to more iterations to find good solutions. There is no clear way to avoid this handicap when comparing networks of varying numbers of APs, but by running the algorithm until we reach our implementation's limit, we provide the best answer we can in each case.

Recall in Chapter III.D.2 we bound our solution space by upper and lower bounds on coverage shortfall and delivered network flow (Figure 8). Similarly, we now use the

lower boundaries of coverage shortfall and delivered network flow as references to create our Pareto frontier (we suppress the upper bound on coverage shortfall because of scale, and the upper bound on delivered network flow because it depends on the numbers of APs). The results are displayed in Figure 32.

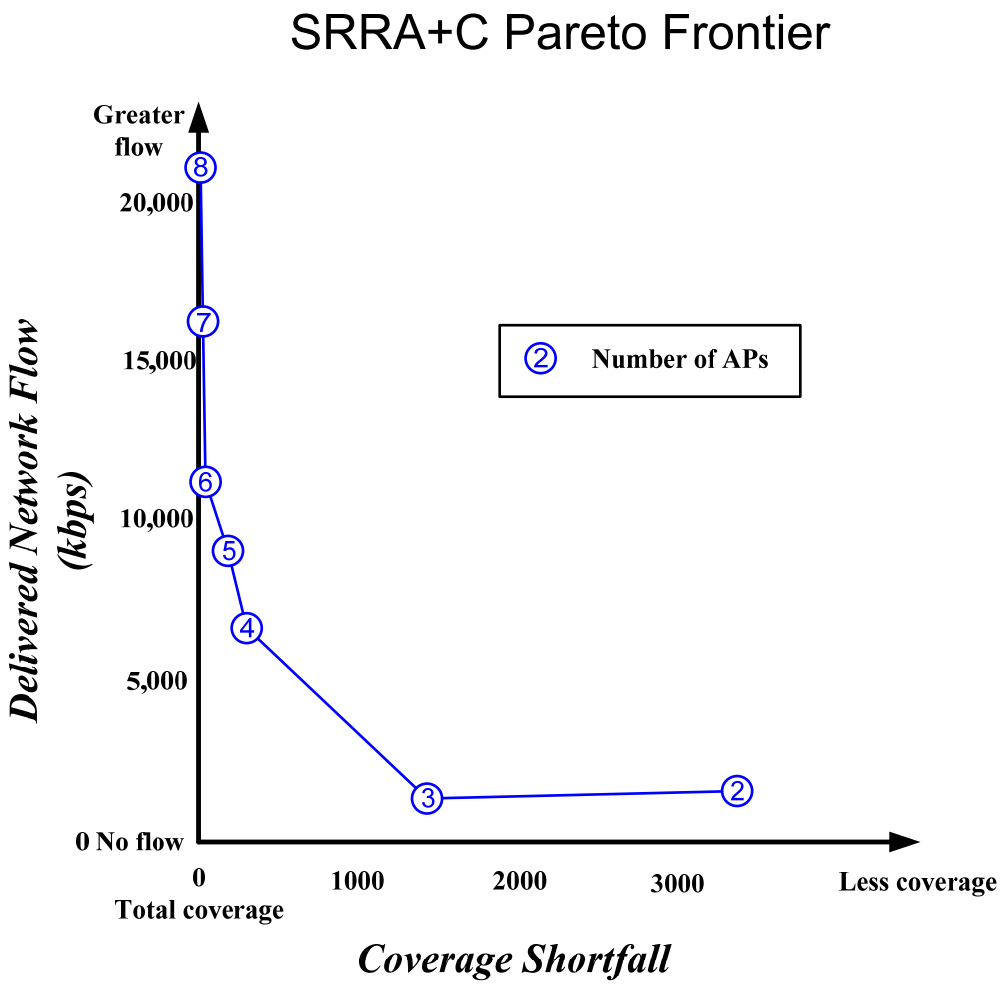


***Figure 32: Example Pareto frontier of SRRA+C***

The Pareto frontier shows a decreasing gain in coverage with additional APs beyond four. The effect of additional APs on delivered network flow is not as straightforward, but clearly more APs provide greater flow. This chart helps answer the questions "How good is our network?" and "How much better can we do?" This information may be of great use to a decision-maker, and the speed of the DIRECT algorithm makes this analysis possible.

## 6. Simple Test Case Analysis

We next use SRRA+C to evaluate simple network scenarios, and present screenshots of our decision support tool to display the solutions. This analysis enables us to compare the output of SRRA+C to the intuitive solution, and to gain a better understanding of what to expect from SRRA+C.

### *a. Square Coverage Region on Dry Lakebed*

Our first three simple tests occur on perfectly flat terrain, such as a dry lakebed, with no obstructions to radio propagation. We begin with defining a simple square coverage region and solving for a five AP network, with the AP node placed directly in the middle of the coverage region. One would expect the optimal solution to place each of the four APs on a distant side or corner of the coverage region. Figure 33 displays the output of the decision support tool for this test.

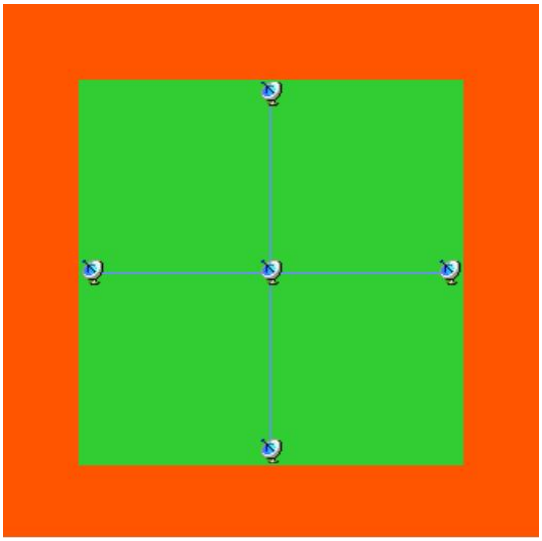

***Figure 33: SRRA+C solution for a square coverage region on flat terrain* [color online]**

As expected, the transmitters are placed on distant sides of the coverage region.

### *b. Corridor Coverage Region on Dry Lakebed*

Next, we examine the same dry lakebed terrain, but with a very long, narrow coverage region (a "corridor") and the HQ node located at one end of the corridor. One would expect three mobile APs to line up within the coverage region and link to their nearest neighbors. Figure 34 displays the output of the decision support tool for this test, with the HQ node fixed at the right side of the coverage region.

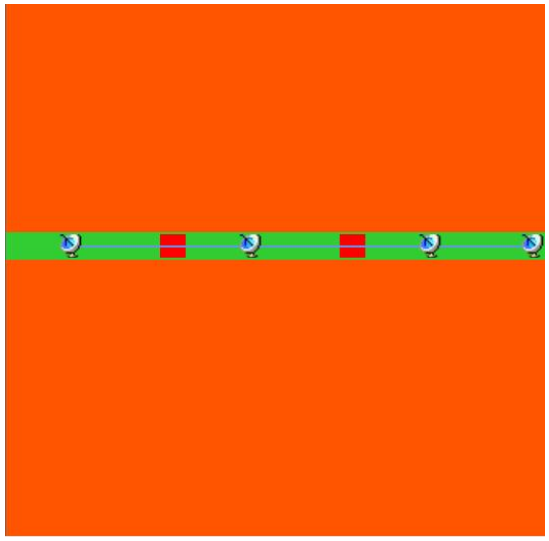

***Figure 34: SRRA+C solution for a corridor coverage region on flat terrain* [color online]**

The formulation performs as expected: each AP links with its neighbor within the defined coverage region.

### *c. Distant Coverage Region on Dry Lakebed*

We next examine a coverage region on a dry lakebed placed at a distance from the HQ node. One would expect three mobile APs to line up and provide a link from the distant coverage region to the HQ node. See Figure 35 for the results of SRRA+C.

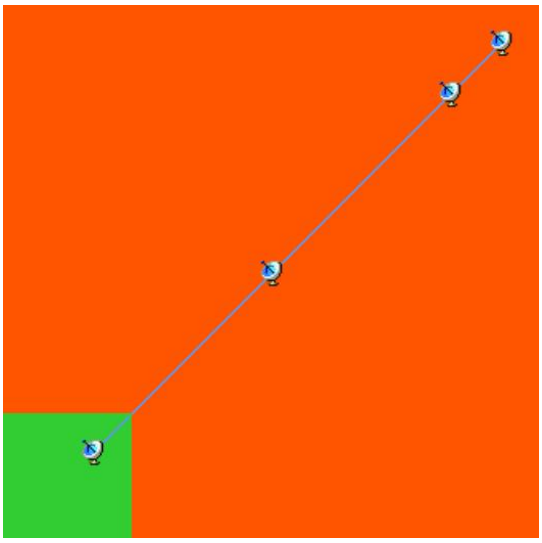

***Figure 35: SRRA+C solution for a distant coverage region on flat terrain* [color online]**

Again, SRRA+C provides the intuitive solution: a straight, linked path from the HQ node to the coverage region.

### *d. Slight Incline*

Our next two test cases occur on real-world terrain of Fort Ord, California. First, we examine the SRRA+C solution when the HQ node is placed near a very shallow hill. One would expect a single mobile AP to be placed on the high side of the hill, enabling optimal radio broadcast downward onto the lower region. See Figure 36 for actual SRRA+C output.

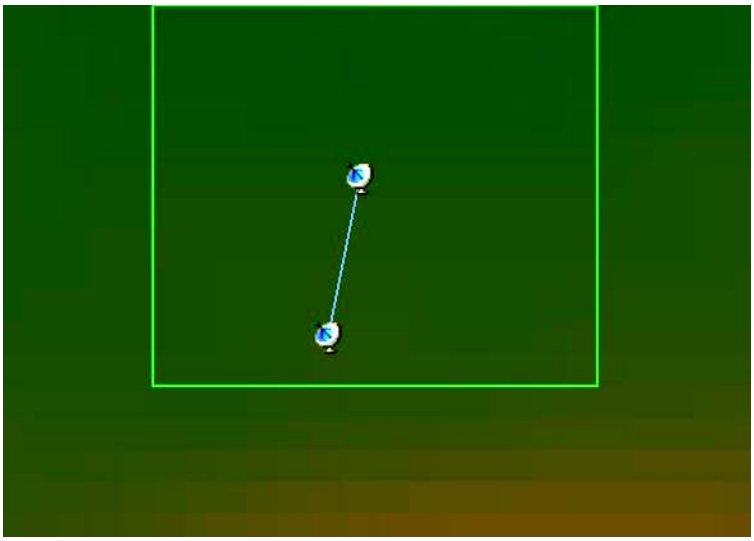

***Figure 36: SRRA+C solution for a coverage region on a very shallow hill* [color online]**

For clarity, coverage shortfall information is suppressed. Though it is difficult to see in the figure, the terrain has a very slight incline rising at the bottom of the figure. As expected, the mobile AP is placed higher than the HQ node on this very slight hill.

### *e. Hill Top*

Our final simple test case places a coverage region on the top of a hill, with the HQ node placed in the center of the coverage region. An intuitive solution is to place a single mobile AP within the coverage region, at a distance from the HQ node but not at too low an elevation, to avoid cutting off the propagation path between the APs. The SRRA+C output is below.

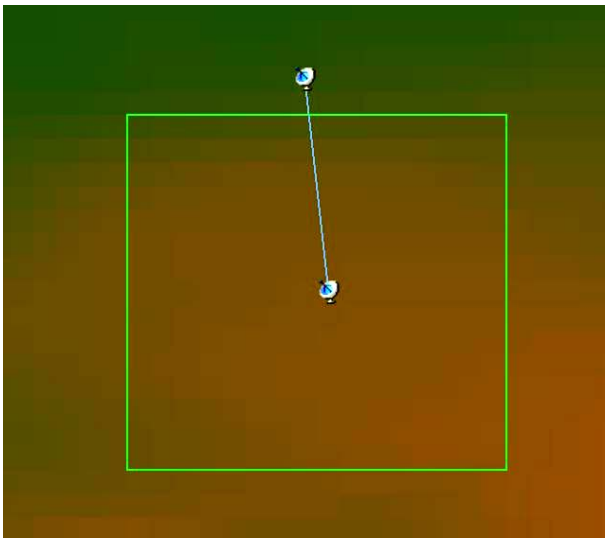

***Figure 37: SRRA+C solution for a coverage region atop a hill* [color online]**

Again, coverage shortfall information is suppressed for clarity. Counter to our intuition, the optimal SRRA+C solution is to place the mobile AP *outside* the coverage region, down the face of the hill. In fact, this solution does seem rational: the lower position (combined with the AP antenna height of two meters) allows the AP to broadcast onto the face of the hill. This maximizes the incident angle between the propagated radio waves and the hill surface, whereas a position nearer the top of the hill would have a much smaller incident angle and hence a potentially smaller received signal strength.

While these small examples certainly do not represent a thorough examination of the behavior of SRRA+C, they do build confidence in its ability to provide sensible solutions.

## C. SOLUTION METHOD PERFORMANCE ANALYSIS

We next compare two methods of determining AP locations to solve the SRRA+C problem. Enumeration calculates every discrete unique solution on the gridded operating region, and so it finds the optimal solution(s) in this discrete space. However, this method is extremely slow as the number of unique solutions increases exponentially with both number of grid locations *UV* and number of APs *n*. If the DIRECT method can find good solutions quickly, it will prove itself much more useful in solving SRRA+C.

We compare these two solution methods by solving for several networks on different types of terrain. Note the runtime required for enumeration prevents us from doing an exhaustive comparison, to include larger operating regions or larger networks.

### 1. Comparison on Theoretical Terrain

We first consider theoretical terrain in the general shape of a volcano, consisting of $UV = 400$ gridded locations (see Figure 38). The terrain is not "smooth," i.e., the face of the volcano resembles a step function. We create a square coverage region and place the HQ node on the lip of the volcano at position (12, 12). We then solve for two, three, and four AP networks using both enumeration and DIRECT. The results of this comparison are tabulated below.

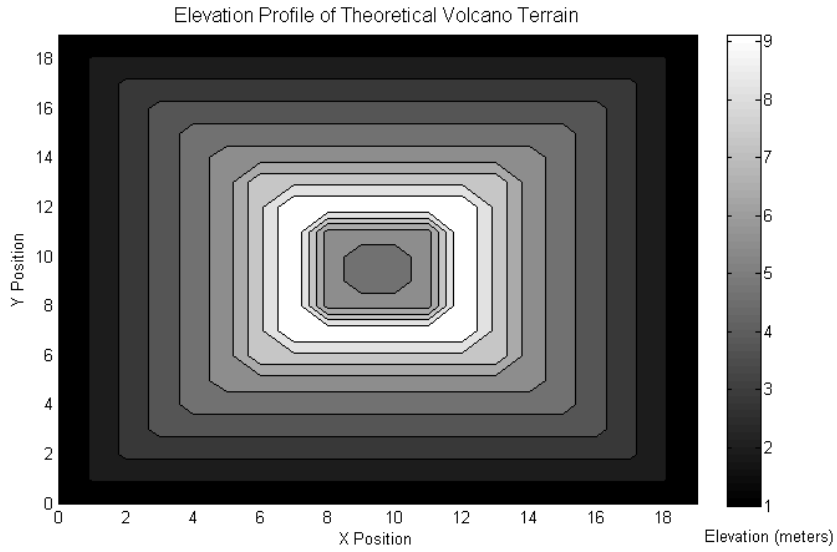


***Figure 38: Contour plot of theoretical "volcano" terrain***

| APs | Number of Unique Solutions | Enumeration Runtime | DIRECT Function Evaluations | DIRECT Runtime | Relative Error of DIRECT Solution |
|---|---|---|---|---|---|
| 2 | 400 | 3 sec, 415 ms | 54 | 179 ms | 14.8% |
| 3 | 79,800 | 17 min, 1 sec, 504 ms | 134 | 595 ms | 4.85% |
| 4 | 10,586,800 | 2 days, 11 hr, 10 min | 406 | 2 sec, 899 ms | 7.69% |
| 5 | 1,050,739,900 | 245 days (*extrapolated*) | 454 | 6 sec, 415 ms | ? |

***Table 3: Results of enumeration and DIRECT on "volcano" terrain***

This test illustrates several important points. First, DIRECT is substantially faster than enumeration, and in this example, provides decent solutions to the SRRA+C problems. We note that the runtime of both solution methods increases roughly linearly with the number of function evaluations. This example also illustrates that the explosive growth of enumeration runtime makes comparison between DIRECT and enumeration infeasible for anything larger than trivial networks. Hence, we are unable to compare DIRECT's results of the five AP network to an enumerated solution.

In this example, DIRECT falls into local optima, and while running the algorithm for more iterations alleviates this, we maintain this example to illustrate the importance of terrain on the SRRA+C solution space and DIRECT's ability to converge to optimality. Using the Point-to-Point analysis mode of our decision support tool on this theoretical

terrain, we have found circumstances where very small adjustments in AP position (less than a discrete grid square) have resulted in calculated link capacities increasing by several orders of magnitude. This ruggedness of the solution space greatly increases the difficulty in finding the optimum. As we see in the next example, actual terrain is less likely to cause these problems.

### 2. Comparison on Actual Terrain

We now conduct a similar comparison using a small area of actual terrain on Fort Ord. The terrain consists of *UV* = 441 gridded locations covering 9.6 acres (see Figure 39). The terrain is much smoother than our hypothetical terrain in the above section, so we expect better performance from DIRECT. We again create a square coverage region and place the HQ node directly in the middle, and solve for networks of two to five APs using enumeration and DIRECT. The results of this comparison are tabulated below.

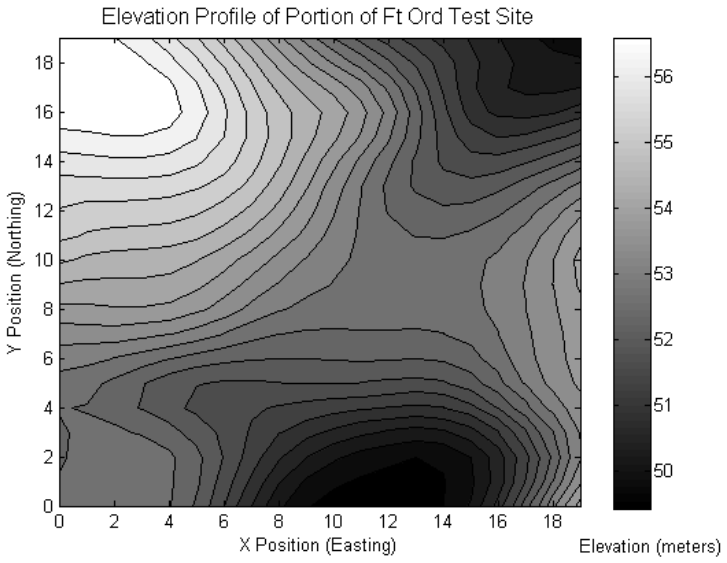


***Figure 39: Contour plot of small section of Ft Ord terrain***

| APs | Number of Unique Solutions | Enumeration Runtime | DIRECT Function Evaluations | DIRECT Runtime | Relative Error of DIRECT Solution |
|---|---|---|---|---|---|
| 2 | 441 | 3 sec, 713 ms | 42 | 554 ms | *-1.3%* |
| 3 | 97,020 | 17 min, 3 sec, 33 ms | 114 | 798 ms | 0.012% |
| 4 | 14,197,260 | 3 days, 26 min, 38 sec, 81ms | 1010 | 9 sec, 379 ms | 2.23% |
| 5 | 1,554,599,970 | ≈ 331 days (*extrapolated)* | 1114 | 16 sec, 954 ms | ? |

***Table 4: Results of enumeration and DIRECT on actual terrain***

Not only is DIRECT much faster than enumeration, but in some cases it finds solutions that are actually *better* than enumeration (indicated by italics and negative relative error values)! This is possible because of the continuous nature of the DIRECT algorithm. DIRECT may place APs literally anywhere within the solution space, whereas the enumeration technique is limited to placing APs at discrete gridded locations.

In fact, as the number of iterations goes to infinity, DIRECT is *guaranteed* to find a solution at least as good as enumeration, and will almost certainly find a better one. Enumeration considers a relatively sparse subset of solutions, whereas DIRECT considers an increasingly dense subset as the number of iterations goes to infinity. Thus, DIRECT is very likely to find a better solution than discrete enumeration.

While this result is exciting, it highlights the disjointedness of comparing continuous and discrete solution techniques. In Chapter V, we recommend comparison of DIRECT with other solution methods.

To further illustrate the usefulness of DIRECT, we return to the example of Chapter IV.B.1. We consider a three AP network on actual Fort Ord terrain consisting of $UV = 2145$ discrete grid locations. Recall we enumerate the 2,299,400 unique solutions in eight hours, 38 minutes, and 55 seconds, and then present the solution objective value as a function of solution rank. We now run DIRECT on this same network and terrain. The results appear in Figure 40.

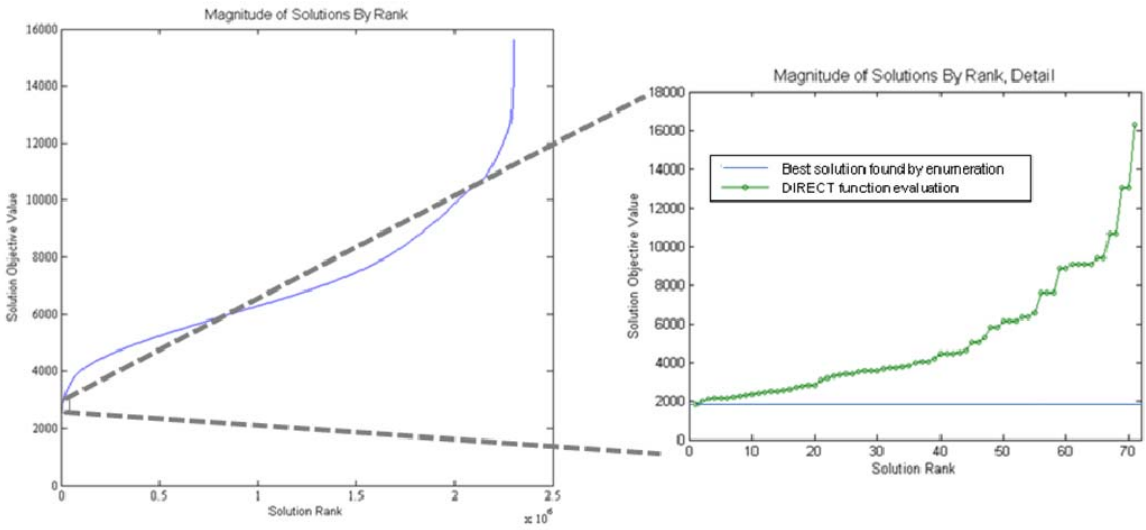


***Figure 40: Comparison of enumeration and DIRECT solution objective values on actual terrain***

The left chart displays the results of enumeration over all unique solutions. The right chart displays an extreme close-up of the minima of the left chart, comparing the results of DIRECT and enumeration. DIRECT not only finds an answer slightly *better* than enumeration, but it does so in only two seconds and 50 milliseconds, requiring a total of just 71 function evaluations.

## D. FIELD TESTING

Our final analysis of SRRA+C with DIRECT is a full network field experiment aboard Fort Ord, conducted in collaboration with the Hastily-Formed Networks (HFN) Research Group at the Naval Postgraduate School. We use the same testing equipment as described in Chapter IV.A. Client coverage is provided via 802.11b/g protocol (2.4 GHz) and the backhaul network is provided via 802.11a protocol (5.8 GHz).

### 1. Operating Region

The operating region for our field experiment is a 45 acre rectangular region aboard Fort Ord. The terrain has moderate foliage, and consists of grass fields, pavement, hard-packed gravel lots, and several buildings. The northern half of the

operating region has low rolling hills, whereas the southern half is generally flat. Elevation ranges from 143 to 226 feet above sea level.

Figure 41 is an overhead image of the operating region (courtesy of Google Maps). Figure 42 is a three-dimensional image of the operating region, viewed from the northwest. Figure 43 is an overhead contour plot of elevations. The scales of Figures 42 and 43 are relative to the position (606462.915 E, 4056897.317 N) UTM zone 10, with an easting (*x* dimension) scale of 8.455 meters and a northing (*y* dimension) scale of 10.447 meters. Figure 44 is a screenshot from our decision support tool of the operating region. Recall our decision support tool represents lower elevations as green, and higher elevations as orange (in black and white, these areas are respectively darker and lighter).


***Figure 41:*** ***Aerial view of Fort Ord operating region*** **[color online] [image courtesy of Google Maps]**


***Figure 42: Elevation profile of operating region, viewed from the northwest***

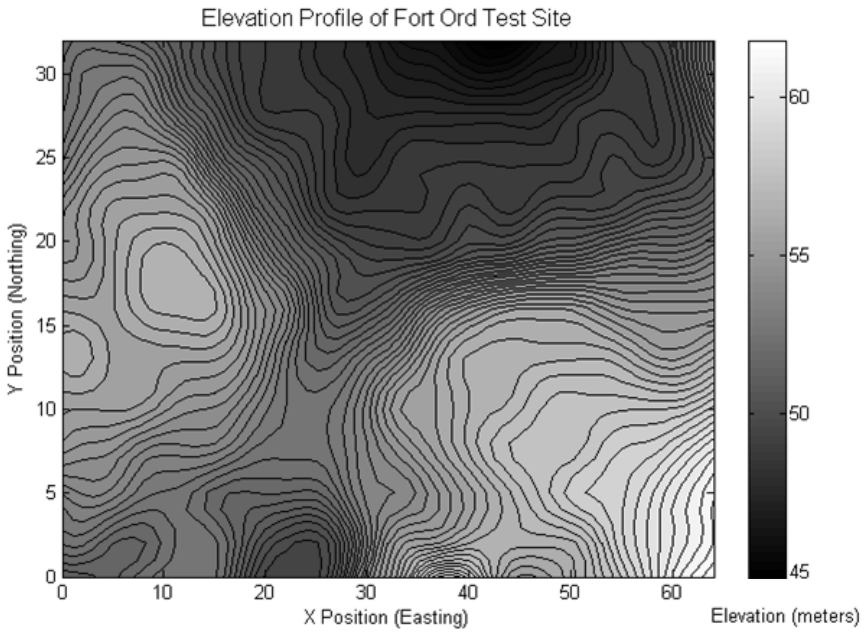


***Figure 43: Contour plot of operating region elevation profile***

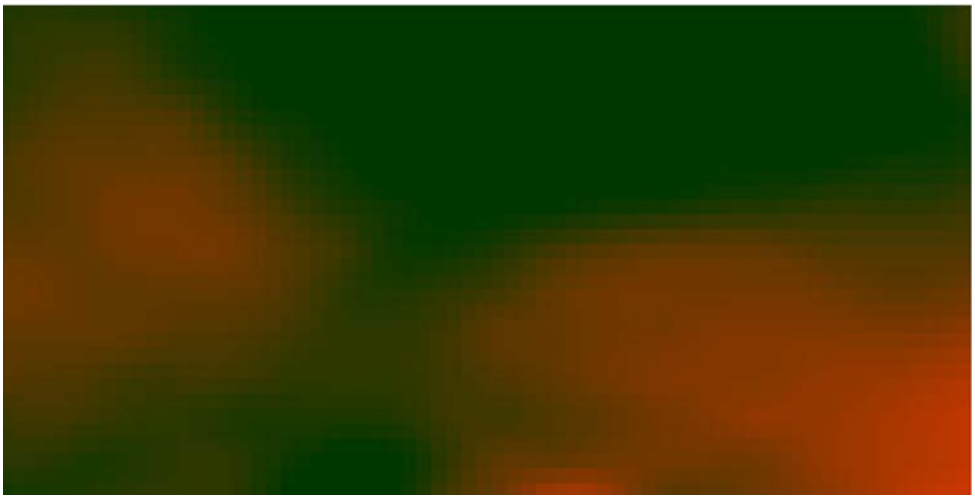

***Figure 44: Screenshot of Fort Ord operating region using decision support tool* [color online]**

We use the Cognio Mobile Spectrum Management system (Cognio, 2005) to obtain an estimate of the background noise level of our operating region. This information (Figure 45) is used as an input to our decision support tool. All of the networks are designed using a weight *w* of one. See Appendix B for a complete list of inputs.

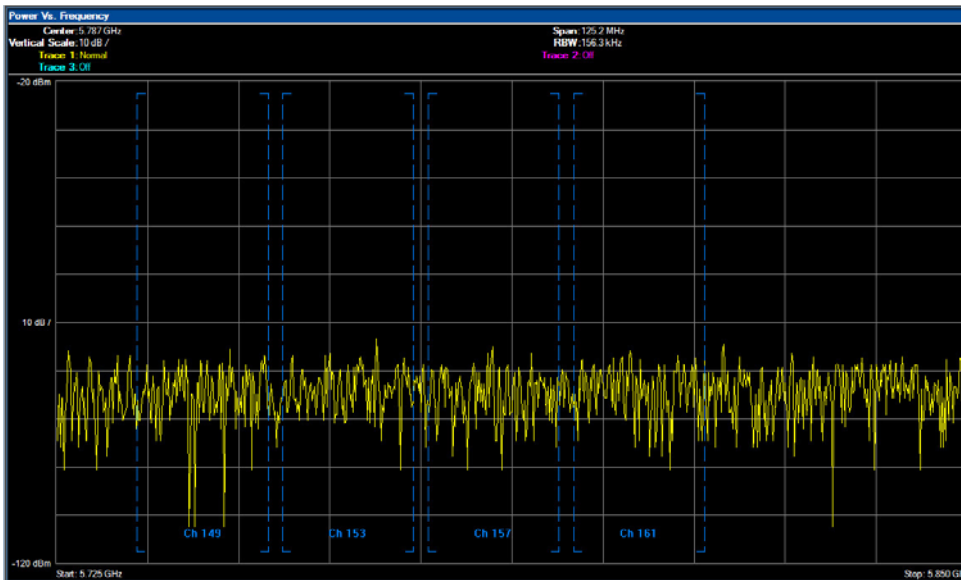


***Figure 45: Screenshot of Cognio Mobile Spectrum tool measuring background noise. Horizontal axis represents frequency in Hz, and vertical axis represents background noise amplitude in dBm* [color online]**

### 2. Scenario Descriptions and SRRA+C with DIRECT Solutions

We create three scenarios requiring the employment of a WMN, and then use our decision support tool to find AP locations and expected throughput for each scenario. Screenshots are used in this section to display AP locations; expected and actual throughput values are detailed in the next section.

#### *a. Distant Coverage Region*

Our first scenario requires the connection of a distant client coverage region with a fixed HQ node, located outside the coverage region. For instance, emergency responders require on-scene network connectivity to their distant command post. Given a total of three APs, an intuitive solution to this problem places one AP at the HQ node, one AP servicing the desired coverage region, and the other AP serving as a bridge between the first two.

We use our decision support tool to define a small coverage region in the southwest corner of the operating region, and place the HQ node in the northeast corner. Our tool solves the associated SRRA+C problem with DIRECT in 452 milliseconds, requiring 70 function evaluations. Figure 46 displays the output, with node labels (recall our decision support tool represents areas with adequate client coverage as bright green, and coverage shortfall as red, or respectively lighter and darker in black and white).

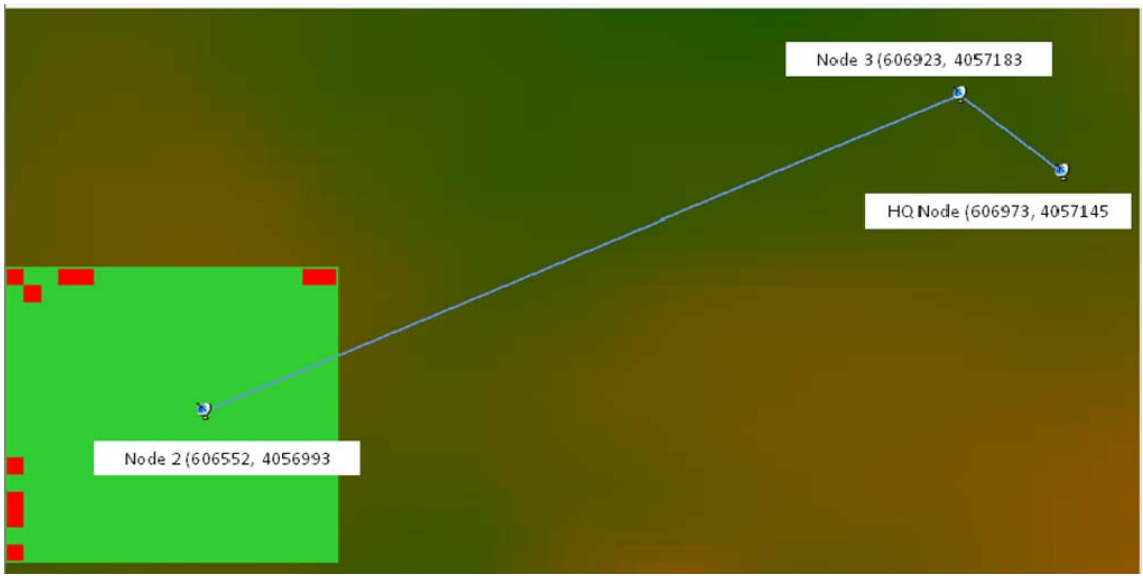


***Figure 46: Solution to distant coverage region scenario* [color online]**

As expected, the third node is positioned as a link between the coverage region and the HQ node. Note that the link between nodes 2 and 3 seems to follow a depression in the terrain, where it is more likely for the radio waves to pass unhindered.

### *b. Coverage Regions of Different Thresholds*

Our second scenario requires the connection of two distant, similarly-sized client coverage regions of different thresholds with a fixed HQ node. Recall from Chapter III a higher coverage threshold value assigns a greater value to the associated coverage region; failure to meet the larger threshold results in a greater penalty (coverage shortfall). This higher threshold can indicate greater importance or greater expected client demand. An example of such a scenario is a disaster relief operation where a distant HQ node is providing Internet connectivity to two separated groups with different numbers of expected users. Given a total of four APs, an intuitive solution to this problem places one AP at the HQ node, one AP servicing the coverage region with the lesser threshold, and the other two APs servicing the coverage region with the greater threshold.

We use our decision support tool to define a low threshold coverage region in the southwest corner of the operating region, and a high threshold coverage region in the northeast corner. We place the HQ node in the southeast corner. Our tool solves the associated SRRA+C problem with DIRECT in 1 second and 437 milliseconds, requiring 162 function evaluations. Figure 47 displays the output, with node labels.

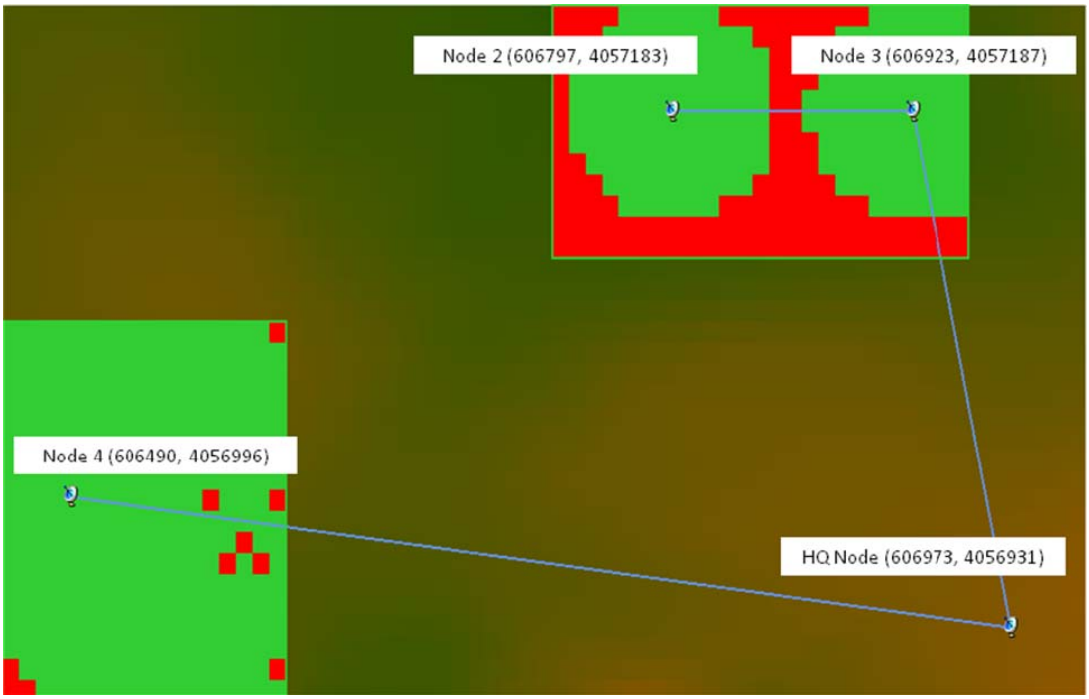


***Figure 47: Solution to differing coverage regions scenario* [color online]**

As expected, the algorithm assigns two nodes to the coverage region with a greater threshold. Note the difference in coverage patterns between the two regions as a result of the different coverage thresholds.

### *c. Large Contiguous Coverage Region*

Our final scenario requires a single large contiguous coverage region. An example of this scenario would be a combat patrol requiring continuous connectivity with the HQ node in order to relay time-sensitive intelligence. Given a total of five APs, an intuitive solution to this problem distributes APs somewhat evenly throughout the coverage region.

We use our decision support tool to define the coverage region as the entire 45 acre operating region. We place the HQ node in the southeast corner. Our tool solves the associated SRRA+C problem with DIRECT in 2 seconds and 538 milliseconds, requiring 166 function evaluations. Figure 48 displays the output, with node labels.

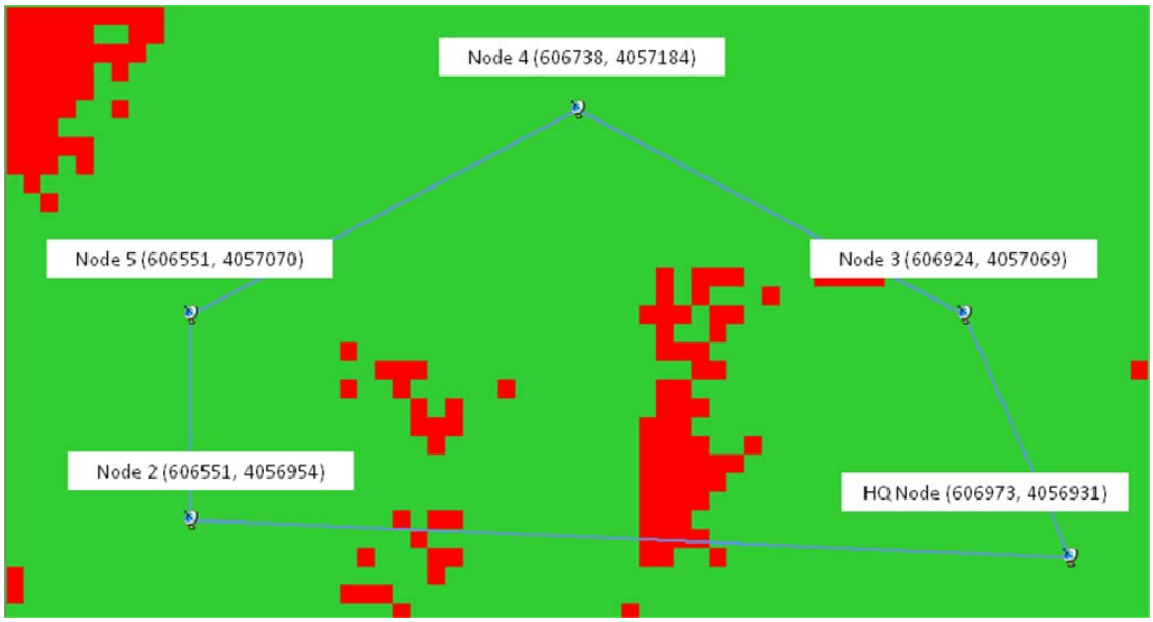


***Figure 48: Solution to large contiguous coverage region scenario* [color online]**

The algorithm creates a ring network topology that does indeed roughly spread the APs throughout the coverage region. Dead spots in client coverage appear in areas of low elevation or behind hills.

### 3. Testing

We create the above topologies using the same Cisco Aironet APs described in Chapter IV.A. We position each AP using a Global Positioning System (GPS) device, and place atop a two meter mast (see Figure 49). When required, we make small adjustments (within five meters) to AP position to avoid broadcasting directly into trees, etc. We use small portable generators and car batteries as power sources. Network throughput is tested using the same methodology described in Chapter IV.A.

***Figure 49: Wireless mesh access point, mast, and portable generator* [color online]**

### 4. Results

For each scenario described in Chapter IV.D.2, we present the throughput values our decision support tool predicts, and the values we actually measure during the field test.

#### *a. Distant Coverage Region*

Our first scenario attempts to connect a distant coverage region to the HQ node via an intermediary node. The table below displays the results. The intermediary node is able to connect to the HQ node, but the node providing client coverage is unable to connect to the intermediary node.

| Node | Predicted Throughput | Actual Throughput |
|---|---|---|
| 2 | 1849 kbps | NONE |
| 3 | 32,265 kbps | 5160-6226 kbps |

***Table 5: Results of distant coverage region scenario***

We believe this lack of connection is due to trees and chain-link fence obstructing the propagation path. This highlights one weakness of the propagation model: recall that TIREM does not account for any vegetation or man-made structures. The left side of Figure 50 shows the view from node 2 to 3, and the right side shows the view from node 3 to 2. A 5.8GHz signal is not very likely to penetrate such foliage.

***Figure 50: View between nodes 2 and 3 of first scenario*** **[color online]**

The actual throughput between node 3 and the HQ node is much less than predicted by the Shannon upper-bound. This observation makes sense when viewed with the results of Figure 26. The distance between node 3 and the HQ is less than 63 meters, and at shorter distances, Figure 26 shows that the Shannon curve is orders of magnitude greater than actual throughput.

### *b. Coverage Regions of Different Thresholds*

The second scenario attempts to connect two separated coverage regions of different thresholds with the HQ node. The table below displays the results. All nodes are able to form a connection to the HQ node, and the observed throughput values are rather close to the predicted Shannon upper-bounds.

| Node | Predicted Throughput | Actual Throughput |
|---|---|---|
| 2 | 2042 kbps | 2867-3277 kbps |
| 3 | 2169 kbps | 2867-3277 kbps |
| 4 | 1353 kbps | 8388-9227 kbps |

***Table 6: Results of scenario with coverage regions of different thresholds***

#### *c. Large Contiguous Coverage Region*

Our final scenario creates a large coverage region over the entire 45 acre operating area. The table below displays the results. Again, each node is able to connect to the HQ node, and actual throughput values are within an order of magnitude of the predictions.

| Node | Predicted Throughput | Actual Throughput |
|---|---|---|
| 2 | 1277 kbps | 6554-7373 kbps |
| 3 | 5295 kbps | 6963 kbps |
| 4 | 3605 kbps | 3686 kbps |
| 5 | 2660 kbps | 3277-3686 kbps |

***Table 7: Results of large contiguous coverage region scenario***

## E. DISCUSSION

Our analysis and field testing of SRRA+C with DIRECT leads us to the following observations:

### 1. TIREM Can Provide Accurate Received Signal Strength Predictions

The results of our point-to-point field test in Chapter IV.A.1 show that with the proper selection of fade margin, TIREM received signal strength predications can be very close to actual observations. This provides evidence that our formulation of coverage shortfall is a valid method of calculating client coverage.

### 2. The Shannon Capacity Formula Can Serve as an Approximation of Expected Throughput

The results of our point-to-point field test in Chapter IV.A.2 show that with an accurate measurement of background noise level and selection of an applicable fade margin, the Shannon capacity formula can approximate link capacity. However, given that the formula is strictly theoretical and is especially impractical at short distances, the expected throughput values should only be used as a relative gauge of network performance.

### 3. The SRRA+C Solution Surface May Have Many Local Optima

Our analyses in Chapters IV.B.1 and IV.B.2 show that the SRRA+C solution surface may be very rugged, with potentially many local optima and peaks of bad solutions. This increases the likelihood that DIRECT will fall into bad local optima, but the algorithm is guaranteed to break out eventually and converge to the global optimum.

### 4. Flow Value Weight *w* Can Serve as a Method of Tuning Network Topology and Flow

Our analysis in Chapter IV.B.1 shows that a larger flow value weight *w* can increase network flow values and decrease the sum of distances between outlying nodes and the HQ node. However, our method of estimating the solution of the SRRA problem (Xiao et al., 2004) can cause minor nonmonotonicities in increasing *w*. We can avoid this by solving the SRRA problem to optimality, but at a greater computational cost.

### 5. SRRA+C Provides Generally Intuitive Network Topologies

The simple test cases we analyze in Chapter IV.B.4 show that SRRA+C can provide seemingly reasonable AP locations. Occasionally, SRRA+C provides solutions that are not intuitively obvious, but in fact do constitute good design.

### 6. DIRECT Can Provide Good Solutions to the SRRA+C Problem Faster Than Enumeration

We show in Chapter IV.D.2 that in practice on actual terrain, DIRECT frequently finds better solutions than discrete enumeration very quickly. DIRECT is guaranteed to find a solution at least as good as enumeration, and every iteration of DIRECT is guaranteed to provide a solution at least as good as the previous.

### 7. Enumeration Has Limited Usefulness as a Method of Comparison to DIRECT

The enumeration method will sample only a subset of discrete unique solutions, whereas the DIRECT method will, as the number of iterations goes to infinity, sample within an arbitrary distance of any point in the solution space. Hence, DIRECT can always provide a solution at least as good as enumeration, and can virtually always provide a better solution. This disjointedness, combined with the exponential growth of the solution space and corresponding enumeration runtimes, implies that comparison of the two methods should be limited to small test cases and gauging how long it takes DIRECT to find an answer at least as good as enumeration.

### 8. SRRA+C with DIRECT Can Quickly Provide Working Network Designs

Our results in Chapter IV.D show that in a realistic field test, SRRA+C solved with DIRECT is capable of providing working network solutions with no guesswork. These results are especially impressive considering the utilized backhaul network transmit power (100 mW) and frequency (5.8 GHz) are not particularly well-suited to outdoor deployments.

### 9. The Underlying Predictive Models Limit the Real-World Accuracy of SRRA+C

The results of the field test show that SRRA+C solutions are only as good as the associated predictions of received signal strength and throughput. The usefulness of TIREM is limited by its inability to account for certain environmental factors such as

vegetation, man-made obstacles and rain. The Shannon capacity formula is strictly designed to provide a theoretical upper-bound of data transmission rates. The modular nature of our formulation allows other predictive models to be used, but greater predictive accuracy may come at a higher computational cost.

### 10. There is No Guarantee Any SRRA+C Solution Will Work in the Real World

Regardless of underlying predictive models, the SRRA+C formulation remains theoretical and provides no guarantee of real world applicability. The formulation does not consider the effects of signal modulation scheme, proprietary traffic routing algorithms, or other device-specific variables.

### 11. We Cannot Provide Certificates that Guarantee the Optimality of Any SRRA+C Solution

As discussed in Chapter III, the DIRECT algorithm is guaranteed to eventually converge to the global optimum. However, we cannot guarantee the optimality of any particular solution during the course of the algorithm, as the distance to the global optimum is never known.

In the hasty applications for which we designed this technique, proof of optimality is not necessary. In fact, unrelenting pursuit of optimality in such a complex problem wastes a very critical resource: time. Rather, good working solutions are required very quickly. As shown, our technique can provide this functionality to decision-makers, as well as quantitative information on the goodness of a solution using the theoretical lower bound and Pareto frontier analyses.

THIS PAGE INTENTIONALLY LEFT BLANK

# V. CONCLUSION AND RECOMMENDATIONS

We conclude by summarizing our key findings, listing potential applications of our work, and making recommendations on how this work can be further extended.

## A. SRRA+C PROVIDES A NUMERIC GAUGE OF NETWORK PERFORMANCE

We developed a method of quantifying WMN performance based on maximizing client coverage, subject to constraints on network flow, quantity and technical characteristics of APs, desired coverage region, and radio propagation over terrain.

We calculate client coverage by summing the positive differences between desired and actual received signal strength at each discrete location within the user-defined coverage region. We estimate received signal strength using either the Hata COST–231 model or TIREM.

We quantify network flow by adopting the SRRA formulation of Xiao et al., (2004). The formulation simultaneously calculates traffic routing and AP transmission resource allocation. We solve the problem using the subgradient method for a given number of iterations to estimate the value of network flow.

We combine client coverage and network flow to produce a numeric value of a given WMN. We show that the flow value weight $w$ can be used to tune network performance and topology to meet certain criteria.

Our SRRA+C formulation does not require information about device-specific characteristics, such as modulation scheme or routing protocol. While this may reduce the predictive capabilities of the formulation, it is not impossible to add consideration of this information. Further, this general approach makes it very easy to quickly model networks of greatly varying devices and capabilities.

## B. SRRA+C WITH DIRECT QUICKLY PROVIDES GOOD WMN DESIGN SOLUTIONS

Using our decision support tool, we compared the performance of enumeration and the DIRECT algorithm in solving the SRRA+C problem. We showed that the DIRECT algorithm can find good solutions much faster than enumeration, and is capable of finding better solutions than enumeration due to DIRECT's continuous nature. DIRECT is guaranteed to find the optimal solution to the SRRA+C problem as the number of iterations goes to infinity, and each iteration of DIRECT provides a solution at least as good as the previous.

To our knowledge, we are the first to use an algorithm with proven global convergence to solve a WMN design problem to maximize client coverage, given constraints on network service, client coverage, quantity and characteristics of AP devices, environmental information, and radio propagation over terrain. Our technique and associated tool runs on a laptop computer and does not require any additional software or solver licenses.

Through several field tests, we showed that the SRRA+C problem solved with DIRECT can quickly provide working WMN topologies in real-world scenarios. This technique requires very little technical expertise and no guesswork.

## C. POTENTIAL APPLICATIONS

Our techniques and associated decision support tool can be used by HA/DR personnel and combat communications planners to quickly design WMNs to support their specific operations. The decision support tool accepts map data in a generic file format that is widely available on the Internet, and can create network topologies for virtually any type of terrain and mesh AP device.

Our decision support tool is written in the same programming language as the U.S. Marine Corps' SPEED communications tool, and could be integrated into that software as a WMN planning module.

The speed of our technique makes it attractive as an engine in automated network design within large combat simulation models. These models must often be run many thousands of iterations, and processing time is at a premium. Our technique could quickly provide realistic command-and-control topologies for these simulations.

## D. RECOMMENDATIONS FOR FUTURE WORK

### 1. Use of More Accurate Radio Communication Models

We note in Chapter IV that the real-world accuracy of our technique is only as good as the underlying predictive models. The modular nature of our formulation allows the use of more accurate radio propagation and network flow models. Future work could investigate such models and compare the benefit of increased accuracy to any additional computational workload. The integration of radio propagation models that consider vegetation and man-made obstructions, and more realistic network flow models that can quickly be solved exactly would be particularly interesting.

### 2. Enable Use of Transmitter Restriction Zones with DIRECT

Our decision support tool allows the user to define *transmitter restriction zones*, where APs are not allowed to be placed. In a real-world scenario, there may be certain areas (such as airfields or roadways) where it is impractical or impossible to place APs, and this functionality would enable such restrictions. Our enumeration technique can utilize this information, but in its current form the DIRECT algorithm cannot, as any such restriction would make the solution space discontinuous.

Future work could leverage this functionality by determining a large but finite penalty for placing APs within a transmitter restriction zone. This would discourage the DIRECT algorithm from choosing locations within these zones, but would not provide a guarantee of such behavior. Another approach may be to create a transformation of the operating region by "cutting out" the transmitter restriction zones and "pasting" the remaining pieces together, creating a continuous solution space usable by DIRECT.

### 3. Method of Automatically Choosing Weight *w*

Flow value weight *w* allows the user to tune the topology and network flow of the optimal SRRA+C solution. Future work could consider methods of automatically choosing this weight, given desired topology or network flow information. Another approach could consider calculating solutions over a range of *w* values, and presenting the results as a Pareto frontier similar to Raisanen and Whitaker (2005) and Chapter IV.B.5.

### 4. Comparison of DIRECT with Other Solution Techniques

We note in Chapter IV that the value of comparing the results of enumeration and DIRECT is limited by their respective discrete and continuous natures. To provide a better understanding of how good DIRECT-provided solutions are (and how quickly they are obtained), future work could consider comparing DIRECT to other sampling algorithms or heuristic approaches.

### 5. Optimizing Flow to All AP Nodes

In the network topologies we consider, we optimize flow from all outlying nodes to a fixed HQ node. In general, this approach creates tree network structures. However, the SRRA formulation of Xiao et al., (2004) allows optimization to any number of destination nodes, and nothing in our technique forbids it. Future research could investigate the effects of destination node selection on network topology.

### 6. Incorporation of Directional Antenna

Our approach assumes all antennae are omnidirectional, and hence does not determine antenna direction. Future research could consider directional antennae, but this will likely come at a heavy computational cost. An additional dimension would need to be associated with each AP to determine antenna direction, and if client and backhaul antennae are facing different directions, a fourth dimension must be added. If both direction and tilt (vertical angle) are considered, more dimensions are required. These

added dimensions would very likely increase the probability of DIRECT falling into a bad local optima, and many more iterations would be required to break out.

### 7. Incorporation of Multiple AP Types

One of our assumptions is that all APs have exactly the same characteristics. However, our formulation and implementation of DIRECT is fully capable of handling multiple AP types. Future work could consider the implications of solving the SRRA+C problem without this assumption.

### 8. Design for Network Resilience

Future research could incorporate the notion of disruptions (accidental or intentional) and the implications on network design (e.g., Grotschel et al., 1995, and Shankar 2008). This would enable a user to determine weak points within a potential network topology and make adjustments accordingly.

### 9. Incorporation of Stochastic Model of Client Demand

Our research assumes client demand is evenly distributed within a defined coverage region. Future research could incorporate a stochastic model of demand, probabilistically distributing demand throughout a coverage region. This would remove the deterministic property of our formulation and require many separate runs to provide statistically meaningful results, but the quick processing times of DIRECT ensures this is feasible.

### 10. Integration of Temporal Information and Mobile Access Points

Future research could consider the implications of coverage regions and associated client demands changing as a function of time, and the mobility of APs to support such regions and demands. For instance, the user could associate required times with coverage regions, or define a track the coverage region would follow over time. Such a formulation would likely require a constraint or include an incentive to minimize the movement of individual APs on each time step or event, to ensure that the movement

of each AP (mounted on a vehicle or unmanned system) is realistic. This research would have clear applicability to the U.S. Marine Corps' Communications-on-the-Move (COTM) concept (Kreisher, 2009).

### 11. Use of Parallel and Multiple Processor Technology

Note the DIRECT algorithm lends itself naturally to parallel and multiple-processor computing. The initial unit hyper-cube of the solution space can be divided into sectors and processed separately, and the best solutions found in each sector can then be compared upon completion. This property of DIRECT makes it particularly appealing in solving large industrial problems with extremely complex and computationally expensive functions, and is utilized in this fashion by He et al., (2004). While our problem is designed to be solved on a laptop computer, use of multi-threading technology could decrease processing time and enable the consideration of very large operating regions and large numbers of APs. Use of our technique on parallel computing systems could be used to design extremely large, city-size WMNs.

# APPENDIX A: LIST OF DECISION SUPPORT TOOL INPUTS AND OUTPUTS

## A. INPUTS

Our decision support tool receives as input the following information:

### 1. Map Data

Map data files are stored in generic text files, and contain the following information:

- Map title
- Horizontal and vertical scales
- Coordinates of lower-left (southwest) corner of map data
- Number of columns and rows in map data
- Minimum and maximum elevation points of map data
- Tab-separated point elevations in meters

### 2. Access Point and Client Device Characteristics

- Number of APs
- AP and client device antenna height (meters)
- AP and client device antenna gain (dBi)
- AP and client device antenna polarization (horizontal or vertical)
- Transmit power from APs to other APs and clients (watts)
- Backhaul and client coverage transmit frequencies (MHz)
- Backhaul network channel bandwidth (Hertz)

### 3. Network Planning Information

- Headquarters node location
- Client coverage regions, defined by coordinates and received signal strength thresholds (dBm)
- Transmitter restriction regions (not considered in this thesis), defined by coordinates
- Network flow weight *w*
- Fade margin
- Direction to optimize network flow (to HQ node, or to all nodes) (not considered in this thesis)

### 4. Environment Information

- Conductivity of earth's surface (S/M)
- Surface humidity ($G/M^3$)
- Relative permittivity of earth (unitless)
- Surface refractivity (N-Units)
- Background noise (dBm)

### 5. Propagation Options

- Inverse square, Hata COST-231, or TIREM
- Terrain sampling rate (for TIREM only)

### 6. Optimization Options

The following options are specific to the Coverage Analysis mode:

- Optimization algorithm (enumeration or DIRECT)

- Number of iterations
- Alternate solutions to save
- SRRA epsilon value
- DIRECT epsilon value

### 7. Point-to-Point Analysis Options

The following options are specific to the Point-to-Point Analysis mode:

- Test Sites 1 and 2 locations
- Connection type (AP to AP, AP to client, or client to AP)

## B. OUTPUTS

### 1. Coverage Analysis Mode Outputs

Our decision support tool produces the following output in Coverage Analysis mode, following optimization of the SRRA+C problem (all results can be saved to a text file):

- Locations of all APs
- Predicted network capacity between all APs
- Coverage shortfall regions and actual received signal strengths
- Coverage shortfall and network flow values
- Overall objective value
- Given number of alternate solutions (including above information for each solution)
- Number of required function evaluations
- Total optimization execution time

### 2. Point-to-Point Analysis Mode Outputs

- Terrain profile
- LOS path and first Fresnel zone
- Propagation type (LOS, diffraction, or troposcatter)
- Free space and total loss (dB)
- Received signal strength (dBm)
- Theoretical throughput (kbps)

# APPENDIX B: LIST OF INPUTS FOR FIELD EXPERIMENT

The following inputs were used during our network field experiment aboard Fort Ord:

## A. ACCESS POINT AND CLIENT DEVICE CHARACTERISTICS

- AP antenna height: 2 meters
- Client antenna height: 1 meter
- AP antenna gain (backhaul): 6 dBi
- AP antenna gain (client): 6 dBi
- Client antenna gain: 3.5 dBi
- AP and client device antenna gain (dBi)
- Antenna polarization (all): horizontal
- Transmit power from APs to other APs and clients: 0.1 watt
- Backhaul transmit frequency: 5800 MHz
- Client coverage transmit frequenc: 2400 MHz
- Backhaul network channel bandwidth: 16.6 MHz

## B. NETWORK PLANNING INFORMATION

- Client coverage thresholds: -87 dBm for all, except northeast coverage region in four node scenario (-75 dBm)
- Network flow weight *w*: 1
- Fade margin: 30 dB

## C. ENVIRONMENT INFORMATION

- Conductivity of earth's surface: 50 S/M

- Surface humidity: 5 $G/M^3$
- Relative permittivity of earth: 25
- Surface refractivity: 300 N-Units
- Background noise: -87 dBm

## D. PROPAGATION OPTIONS

- TIREM
- Terrain sampling rate: 1 per map unit

## E. OPTIMIZATION OPTIONS

The following options are specific to the Coverage Analysis mode:

- Optimization algorithm: DIRECT
- Number of iterations: 5-10
- SRRA epsilon value: 0.01
- DIRECT epsilon value: 0.0001

# LIST OF REFERENCES


Alion Science and Technology Corporation (2007). "TIREM Details." Retrieved on August 22, 2008 from http://www.alionscience.com/index.cfm?fuseaction=Products.viewpage&productid=19&pageid=34.

Analytical Graphics, Inc. (2009). "TIREM for STK/Communications and STK/Radar." Retrieved on April 4, 2009 from http://www.agi.com/products/desktopApp/stkFamily/modules/extensions/tirem/.

Allen, S.M., Hurley S., Whitaker, R.M. (2002). "Automated Decision Technology for Network Design in Cellular Communication Systems." Proceedings of the 35th Annual Hawaii International Conference on System Sciences, Volume 3, pp. 80-87.

Amaldi, E., Capone, A., Cesana, M., Filippini, I., Malucelli, F. (2008). "Optimization Models and Methods for Planning Wireless Mesh Networks." Computer Networks, Volume 52, Issue 11, pp. 2159–2171.

Balanis, C. (2005). "Antenna Theory: Analysis and Design." Wiley-Interscience, U.S.A.

Beljadid, A., Hafid, A., Gendreau, M. (2007). "Optimal Design of Wireless Mesh Networks." 2007 IEEE Global Telecommunications Conference, pp. 4840–4845.

Bertsekas, D. (1999). "Nonlinear Programming." Athena Scientific, Belmont, Massachusetts.

Bhatnagar, S., Ganguly, S., Izmailov, R. (2006). "Design of IEEE 802.16-based Multi-Hop Wireless Backhaul Networks." Proceedings of the 1st International Conference on Access Networks, ACM International Conference Proceeding Series, Volume 267, Article Number 5.

Brown, G., Carlyle, M., Salmeron, J., Wood, K. (2006). "Defending Critical Infrastructure." Interfaces, Volume 36, Number 6, pp. 530–544.

Calegari, P., Guidec, F., Kuonen, P., Nielsen F. (2001). "Combinatorial Optimization Algorithms For Radio Network Planning." Theoretical Computer Science, Volume 263, pp. 235–245.

Cisco Systems, Inc. (2009). "Cisco 100 Series Lightweight Access Points Data Sheet." Retrieved on April 16, 2009 from http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps6306/ps6315/product_data_sheet0900aecd8025708a_ps6306_Products_Data_Sheet.html.

Cognio, Inc. (2005). "Cognio Announces First Laptop-Based Wi-Fi Spectrum Analyzer." Retrieved on April 27, 2009 from http://www.torwug.org/Articles%5Csubmitted%5CCognio%5CPDF%5Cspectrum%20analyzer%20press%20release.pdf.

"A Concept for Enhanced Company Operations" (2008). United States Marine Corps (USMC).

Corne, D.W., Knowles, J.D., Oates M.J. (2000). "The Pareto Envelope-Based Selection Algorithm for Multiobjective Optimization." Proceedings of the Sixth International Conference on Parallel Problem Solving from Nature, pp. 839–848.

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T. (2000). "A Fast Elitist Nondominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II." Lecture Notes in Computer Science, Volume 1917, pp. 848–849.

European Cooperation in the Field of Scientific and Technical Research (COST) (1999). "Digital Mobile Radio Towards Future Generation Systems: COST 231 Final Report."

Eppink, D., Kuebler, W. (1994). "TIREM/SEM Handbook." Electromagnetic Compatibility Analysis Center, Department of Defense.

Finkel, D. (2003). "DIRECT Optimization Algorithm User Guide." Center for Research in Scientific Computation.

Gonzalez, R.C., Woods, R.E. (2007). "Digital Image Processing." Prentice Hall, U.S.A.

Google Maps (2009). Retrieved on April 2, 2009 from http://maps.google.com/.

Grotschel, M., Monma, C.L., Stoer, M. (1995). "Design of Survivable Networks." Network Models, M.O. Ball, T.L. Magnanti, C.L. Monma, and G.L. Nemhauser, Eds. Handbooks in Operations Research and Management Science, Volume 7. Elsevier North-Holland, pp. 617–669.

He, J., Verstak, A., Watson, L., Stinson, C., Ramakrishnan, N., Shaffer, C., Rappaport, T., Anderson, C., Bae, K., Jiang, J., Tranter, W. (2004). "Globally Optimal Transmitter Placement for Indoor Wireless Mesh Networks." IEEE Transactions on Wireless Communications, Volume 3, Number 6, pp. 1906–1911.

Horst, R., Hoang, T. (1996). "Global Optimization." Springer, U.S.A.

Jones, D.R., Perttunen, C.D., Stuckman, B.E. (1993). "Lipschitzian Optimization Without the Lipschitz Constant." Journal of Optimization Theory and Applications, Volume 79, Number 1, pp. 157–181.

Kreisher, O. (2009). "Fast & Flexible: Marine Corps looks for command and control on the go." Seapower, February 2009, pp. 34–36.

Longley, A.G., Rice, P.L. (1968). "Prediction of Tropospheric Radio Transmission Loss Over Irregular Terrain. A Computer Method-1968." Institute for Telecommunications Sciences, Boulder, CO.

Luss, H., Gupta, S.K. (1975). "Allocation of Effort Resources Among Competing Activities." Operations Research, Volume 23, Number 2, pp. 360–366.

MapMart (2009). "MapMart Global Mapping Solutions." Retrieved on April 2, 2009 from http://www.mapmart.com/.

Microsoft Corporation (2009). "Microsoft Visual Studio 2008 Professional Edition." Retrieved on February 26, 2009 from http://www.microsoft.com/visualstudio/en-us/default.mspx.

Motorola, Inc. (2007). "MeshPlanner: Design Outdoor Mesh Networks." Retrieved on August 22, 2008 from http://www.motorola.com/Business/US-EN/Business+Product+and+Services/Wireless+Broadband+Networks/Mesh+Networks/Mesh+Tools/MeshPlanner_New_US-EN.

National Telecommunications and Information Administration (2008). "Irregular Terrain Model (ITM) (Longley-Rice)." Retrieved on September 1, 2008 from http://flattop.its. bldrdoc.gov/itm.html.

Olexa, R. (2005). "Implementing 802.11, 802.16, and 802.20 Wireless Networks: Planning, Troubleshooting and Operations." Elsevier, Burlington, MA.

Raisanen, L., Whitaker, R. (2005). "Comparison and Evaluation of Multiple Objective Genetic Algorithms for the Antenna Placement Problem." Mobile Networks and Applications, Volume 10, pp.79–88.

Shankar, A. (2008). "Optimal Jammer Placement to Interdict Wireless Network Services." Thesis, M.S., Operations Research. Monterey, CA: Naval Postgraduate School.

Shannon, C. (1949). "Communication in the Presence of Noise." Proceedings of the IRE, Volume 37, pp. 10–21.

Sharkey, J. (2008). "Automated Radio Network Design Using Ant Colony Optimization." Thesis, M.S., Computer Science. Bozeman, MT: Montana State University.

Shubert, B. (1972). "A Sequential Method Seeking the Global Maximum of a Function." SIAM Journal on Numerical Analysis, Volume 9, pp. 379–388.

Shyy, D.J. (2008). Principal Communications Engineer, The Mitre Corporation. Personal communication.

Siewert, S. (2005). "Big iron lessons, Part 2: Reliability and availability: What's the difference?" IBM. Retrieved on March 15, 2009 from http://www.ibm.com/developerworks/power/library/pa-bigiron2/index.html.

"Statement of Requirements for Public Safety, Wireless Communications & Interoperability" (2006). Department of Homeland Security.

"Systems Planning, Engineering, and Evaluation Device (SPEED)" (2008). United States Marine Corps Systems Command (MARCORSYSCOM). Retrieved on October 18, 2008 from https://www.marcorsyscom.usmc.mil/sites/cins/CNS/Tactical%20Networks/pst/SPEED(2-19-08).pdf.

Valenzuela, C.L. (2002). "A Simple Evolutionary Algorithm for Multi-Objective Optimisation (SEAMO)." IEEE Congress on Evolutionary Computation, pp. 717–722.

Vieira, A., Stoneback, D. (2008). "Wireless Coverage Analysis and Prediction: Getting it Right." Retrieved on August 24, 2008 from http://www.blueclipper.com/scte/disk /Papers/CD-ROM%20Papers/Stoneback_TP.pdf.

Xiao, L., Johansson, M., Boyd, S. (2004). "Simultaneous Routing and Resource Allocation Via Dual Decomposition." IEEE Transactions on Communications, Volume 52, pp. 1136–1144.

Zhang, Y., Luo, J., Hu, H., (2006). "Wireless Mesh Networking: Architectures, Protocols, and Standards." Auerbach Publications, New York, NY.

Zitzler, E., Laumanns, M., Thiele. L. (2001). "SPEA2: Improving the Strength Pareto Evolutionary Algorithm." Computer Engineering and Networks Laboratory (TIK), Technical Report 103.

# INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Commanding General, Training and Education Command
   MCCDC, Code C46
   Quantico, Virginia

4. Director, Marine Corps Research Center
   MCCDC, Code C40RC
   Quantico, Virginia

5. Marine Corps Tactical Systems Support Activity (Attn: Operations Officer)
   Camp Pendleton, California

6. Director, Operations Analysis Division
   MCCDC, Code C19
   Quantico, Virginia

7. Marine Corps Representative
   Naval Postgraduate School
   Monterey, California

8. MCCD OAD Liaison to Operations Research Department
   Naval Postgraduate School
   Monterey, California

9. Manager, Systems Engineering & Analysis Business Area
   National Infrastructure Simulation and Analysis Center
   Sandia National Laboratories
   Albuquerque, New Mexico

10. Marine Corps Tactical Systems Support Activity
    SPEED/JNMS TSO
    Camp Pendleton, California

11. Marine Corps Systems Command
    PG–12, Tactical Networks
    Quantico, Virginia

12. J6, Joint Spectrum Center
Annapolis, Maryland

13. Program Manager, RF Engineering Programs Division
Alion Science and Technology
Annapolis Junction, Maryland

14. Deputy Director, Command, Control and Interoperability Division
Department of Homeland Security Science & Technology Directorate
Washington, D.C.

15. Director, Hastily-Formed Networks Research Group
Naval Postgraduate School
Monterey, California

16. Patent Attorney
Code OOC, Naval Postgraduate School
Monterey, California